MARKOW-ENTSCHEIDUNGSMODELLE MIT NICHTKONVERGENTEN GESAMTKOSTEN

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der
Hohen Mathem.-Naturw. Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
zu
BONN

vorgelegt von
HANS-HEINRICH BODEWIG
aus
WEEZE

BONN 1985

MARKOW-ENTSCHEIDUNGSMODELLE MIT NICHTKONVERGENTEN GESAMTKOSTEN

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der
Hohen Mathem.-Naturw. Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
zu
BONN

vorgelegt von
HANS-HEINRICH BODEWIG
aus
WEEZE

BONN 1985

Angefertigt mit Genehmigung der Mathematisch-Naturw. Fakultät der

Universität Bonn

1. Referent: Professor Dr. M. Schäl
2. Referent: Professor Dr. W. Vogel

Tag der mündlichen Prüfung: 19. April 1985

Institut für Angewandte Mathematik
Universität Bonn
Abteilung für Wahrscheinlichkeitstheorie
und
Mathematische Statistik
Direktor: Dr. W. Vogel

Druck: Rheinische Friedrich-Wilhems-Universität Bonn

Vorwort

Die vorliegende Arbeit wurde angeregt durch das Buch "How To Gamble If You Must - Inequalities for Stochastic Processes" von L.E. Dubins und L.J. Savage (1965), sowie Arbeiten aus der Dynamischen Optimierung. Eine Einführung in die Theorie der Dynamischen Optimierung erhielt ich nicht zuletzt durch eine Vorlesung von Professor Dr. M. Schäl. Herrn Professor Schäl, der mich auf das Buch von Dubins und Savage aufmerksam machte, danke ich für zahlreiche Gespräche, Hinweise und Korrekturen.

Danken möchte ich für die Möglichkeit, diese Arbeit im Rahmen des Sonderforschungsbereichs 72 der Deutschen Forschungsgemeinschaft zu verfassen. Zahlreiche Kontakte und Gespräche mit Mitgliedern und Gästen waren mir durch den Sonderforschungsbereich möglich. Inbesondere erhielt ich aus Gesprächen mit Professor W. D. Sudderth, Minneapolis, zahlreiche Anregungen.

Die vorliegende Arbeit wurde mit Hilfe des Textverarbeitungssystems TEX geschrieben. Den für die Implementierung und Wartung dieses Systems Verantwortlichen danke ich für vielfältige Unterstützung und die bereitwillige Hilfe bei auftretenden Fragen.

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41546505_0028

INHALT

I. Vorbemerkungen

Einleitung 7

1. Modellbeschreibung 11
2. Beziehungen zu anderen Modellen 18
3. Die Nutzenfunktion 20
   a. Fast sicher endliche Stoppzeiten 24
   b. Endliche Stoppzeiten 25
   c. Endlich additive Maße 25
4. Randomisierte Pläne 26

II. Eigenschaften optimaler Pläne

5. Optimalitätsbedingungen 33
6. Egalisierende Funktionen 43
7. Überall optimale Pläne 52
8. Werterhaltende Pläne 56

III. Existenz optimaler Pläne

9. Das verlaßbare Modell 62
10. Stationäre optimale Pläne 67

IV. Optimale Markow-Pläne

11. Egalisierende Markow-Pläne 69
12. Existenz ε-optimaler Markow-Pläne im Modell von Dubins und Savage 71
13. Existenz eines ε-optimalen Markow-Plan im $\mathcal{MEM}$ 79
14. Ein Beispiel 82

V. Anhang

15. Fatou Ungleichungen 86
   a. Fast sicher endliche Stoppzeiten 86
   b. Endliche Stoppzeiten 87
   c. Endlich additive Maße 89
16. Symbole 91
17. Bibliographie 92

1 Einleitung

Zeitabhängige Prozesse sind seit langem Gegenstand der mathematischen Betrachtung. So knüpften sich erste mathematische Überlegungen an die Beschreibung einer Folge von Münzwürfen. Dabei wurden Fragen nach der mittleren Anzahl von Erfolgen untersucht. An diese Überlegungen schlossen sich Fragen nach der Steuerbarkeit solcher Prozesse an, zum Beispiel durch die Wahl des Einsatzes. Modelle zu diesen Fragestellungen wurden von R. Bellman (1957) und L.E. Dubins und L.J. Savage (1965) entwickelt. Insbesondere das Modell von R. Bellman fand Anwendung in Lagerhaltung, Medizin, Ökonomie und Physik.

Bei den Untersuchungen gesteuerter Modelle stehen folgende Fragen im Vordergrund:

1. Was sind unter guten Steuerungen?
2. Lassen sich die Größen des Modells berechnen?
3. Welche Eigenschaften haben gute Steuerungen?

Als gute Steuerungen werden sinnvollerweise solche angesehen, die eine vorgegebene Nutzenfunktion maximieren. In dieser Nutzenfunktion, weniger in sonstigen Modellkomponenten, unterscheiden sich die verschiedenen Modelle. So betrachtet die herkömmliche dynamische Optimierung, angeregt durch R. Bellman, die Summe von Einschrittauszahlungen als Indikator für den Nutzen einer Steuerung. Die Nutzenfunktion von Dubins und Savage bewertet dagegen den Nutzen der durchlaufenen Zustände, ohne Zwischengewinne zu berücksichtigen. Beide Modelle haben ihre Vorzüge. So zeigt das Modell von Dubins und Savage größere Verwandtschaft zu den Überlegungen über den Münzwurf, da es auch die Beschreibung von oszillierenden Prozessen möglich macht, was im Modell von R. Bellman nicht vorgesehen ist.

Die zweite Frage läßt sich leichter beantworten, wenn man die Menge der Steuerungen ohne Gewinnverlust einschränken kann. Dabei beschränkt man sich meist auf Steuerungen, die die Vergangenheit nicht berücksichtigen. Solche Steuerungen lassen oft eine einfache Berechnung des maximalen Nutzens zu. Wegen ihrer Unabhängigkeit von der Vergangenheit sind sie außerdem leichter in der Praxis anzuwenden.

In dieser Arbeit wird die Existenz guter gedächtnisloser Steuerungen untersucht. Es wird gezeigt, daß bei endlichem Zustandsraum solche Steuerungen existieren. Das heißt aber auch, daß man bei der Behandlung des hier vorgestellten Modells nur von der Vergangenheit abhängige Steuerungen betrachten muß. Bei diesen Überlegungen wird eine Nutzenfunktion zu Grunde gelegt, die, ähnlich der von R. Bellman, Einschrittauszahlungen berücksichtigt. Außerdem ist es möglich, die Nutzenfunktion von Dubins und Savage aus ihr abzuleiten. Das hier vorgestellte Modell vereinigt damit die Vorzüge beider Modelle.

Im ersten Paragraph wird das Optimierungsmodell vorgestellt. Die Verbindungen zwischen den verschiedenen Modellen - dem Modell von R. Bellman, dem von Dubins und Savage und dem Modell dieser Arbeit - werden im zweiten Paragraph dargelegt. Die Darstellung der in dieser Arbeit verwendeten Nutzenfunktion wird anhand von Sätzen von R. Chen (1976) und Engelbert-Engelbert (1979) im dritten Paragraph diskutiert. Dabei zeigt sich, daß unter leichten Voraussetzungen verschiedene Darstellungen einer Nutzenfunktion möglich sind. Diese Voraussetzungen sind schwächer als die, die üblicherweise für dynamische Optimierungsmodelle gemacht werden.

In einigen Modellen der Statistik konnte durch Berücksichtigung von randomisierten Entscheidungen die Güte von Tests verbessert werden. Deshalb soll im 4. Paragraph untersucht werden, ob in dem hier betrachteten Modell durch den Einsatz randomisierte Entscheidungen bessere Ergebnisse erzielt werden können. Dies ist jedoch nicht der Fall. Es kann nämlich gezeigt werden,daß zu einer randomisierten Steuerung immer eine deterministische mit fast gleich guter Auszahlung angegeben werden kann. Dieses Ergebnis wird mit Hilfe einer Idee von D. Blackwell (1964) gezeigt. Es verallgemeinert bekannte Ergebnisse für postive, negative und diskontierte Modelle.

In vielen Fällen kann gezeigt werden, daß gedächtnislose, randomisierte Steuerungen deterministische Steuerungen mit Gedächtnis ersetzen können. Dies ist oft ein wichtiges Hilfsmittel bei der Suche nach gedächtnislosen deterministischen Steuerungen. An Hand eines Beispiels wird gezeigt, daß diese Technik für das Modell von Dubins und Savage und damit für das hier untersuchte Modell nicht ohne weiteres anwendbar ist.

Im 5. Paragraph werden Eigenschaften der maximalen mittleren Auszahlungsfunktion untersucht. Aus diesen Eigenschaften folgt, daß gute Pläne durch die zwei Eigenschaften 'werterhaltend' und 'egalisierend' charakterisiert werden können. Die Eigenschaft 'egalisierend' wird in Paragraph 6 behandelt. Dabei zeigt sich, daß ein Plan genau dann 'egalisierend' ist, wenn der durch den Plan gesteuerte Prozeß unendlich oft Zustände durchläuft, die keinen zukünftigen Gewinn erwarten lassen. Eine ähnliche Charakterisierung wird für die werterhaltenden Pläne angegeben. Dazu wird ein 'dualer' Prozeß definiert. Mit Hilfe dieses Prozesses wird gezeigt, daß ein Plan werterhalten ist, wenn die Differenz der maximalen, mittleren Auszahlungsfunktionen des ursprünglichen und des dualen Modells im Prozeßverlauf nicht sehr schwankt. Diese Eigenschaft ist in gewisser Weise für einen 'werterhaltenden' Prozeß notwendig.

Zur Herleitung der Ergebnisse in Paragraph 8 wird untersucht, ob eine Steuerung nur im Mittel über alle Pfade gut ist oder ob sie unter Berücksichtigung der jeweils schon beobachteten Vergangenheit hohe Gewinne erwarten läßt. Diese Eigenschaft ist nicht nur von mathematischem Interesse, sie garantiert auch, daß einmal gewählte Pläne während des Prozeßablaufs nicht durch andere ersetzt werden müssen.

Im dritten Kapitel wird die Existenz guter stationärer Pläne untersucht. Es wird gezeigt, daß unter der Annahme, daß der Prozeß zu einem beliebigen Zeitpunkt gestoppt werden kann, ε-optimale stationäre Pläne existieren. Dieses Ergebnis ist eine Verallgemeinerung eines Satzes von W.D. Sudderth (1969). Mit diesem Resultat wird im 10. Paragraph die Existenz eines guten stationären Plans bei endlichem Zustands- und Aktionenraum gezeigt. Eine schwächere Form dieses Satzes wurde von Dubins und Savage (1965) gezeigt.

Beispiele zeigen, daß sich die Ergebnisse des 10. Paragraphen nicht unter allgemeineren Voraussetzungen, wie zum Beispiel abzählbarem Zustandsraum, zeigen lassen, will man nicht die allgemeine Struktur dieses Modells durch Konvergenzannahmen an die Nutzenfunktion aufheben. Deshalb wird im vierten Kapitel die Frage untersucht, unter welchen Bedingungen gute Markow-Pläne existieren. Markow-Pläne sind vom gerade beobachteten Zustand abhängig und von der Zeit, die seit Beginn des Prozesses verstrichen

ist. Im 11. Paragraph wird gezeigt, daß immer ein egalisierender Markow-Plan existiert. Daraus folgt, daß in den Fällen, bei denen man sich auf die werterhaltenden Pläne einschränken kann, optimale Markow-Pläne existieren.

Im 12. Paragraph wird bewiesen, daß es bei endlichem Zustandsraum möglich ist, eine vorgegebene Menge von Zuständen mit einer Markow-Steuerung unendlich oft mit hoher Wahrscheinlichkeit zu durchlaufen. Aus diesem Ergebnis wird im 13. Paragraph die Existenz guter Markow-Pläne in dem hier behandelten Modell bei endlichem Zustandsraum hergeleitet. Ob und wieweit dieses Ergebnis verallgemeinert werden kann, ist zur Zeit eine offene Frage. Durch ein Beispiel im 14. Paragraph wird aber gezeigt, daß die Existenz guter Markow-Pläne bei abzählbarem Zustandsraum nicht gesichert ist, wenn das Bewegungsgesetz durch ein endlich additives, atomloses (diffuses) Maß beschrieben wird. Mit solchen Maßen arbeiten Dubins und Savage in ihrem Buch "How to Gamble If You Must". Eine Anwendung eines Satzes aus dem 12. Paragraph zeigt, daß in diesem Beispiel entscheidend eingeht, daß das Bewegungsgesetz durch ein diffuses Maß beschrieben wird.

Abschließend werden in einem Anhang einige oft benötigte Sätze in der hier benötigten Form dargestellt. Es handelt sich dabei um sogenannte Fatou-Sätze, die die Vertauschbarkeit von Integral und Limes sichern. Außerdem werden Sätze von R. Chen (1976) für die hier benötigten endlichen Stoppzeiten formuliert und bewiesen.

1. Modellbeschreibung.

Grundlage Dynamischer Optimierungsmodelle sind Systeme, die zu verschiedenen Zeitpunkten beobachtet werden. Durch eine Aktion $a \in \mathcal{A}$ gesteuert, bewegt sich das System aus einem Zustand x mit einer bekannten Wahrscheinlichkeit $\rho(x,a)$ in einen neuen Zustand x'. Dabei wird ein möglicherweise negativer Gewinn $r(x,a,x')$ erwirtschaftet. Durchläuft das System eine Folge von Zuständen und Aktionen $h = (x_0, a_1, x_1 ...)$ bis zum Zeitpunkt n, so beträgt der erwirtschaftete Gewinn $u_n(h) := \sum_{t=0}^{n-1} r(x_t, a_{t+1}, x_{t+1})$. Ziel der dynamischen Optimierungsmodelle ist es zu untersuchen, wie die zu erwartende Auszahlung durch geeignete Wahl der Aktionen maximiert werden kann.

Bei den meisten Anwendungen geht man davon aus, daß im Zustand x nicht jede beliebige Aktion gewählt werden kann. Deshalb wird in dieser Arbeit angenommen, daß die Aktion a in einer Menge $\Gamma(x) \subset \mathcal{A}$ von 'zulässigen' Aktionen liegt. Die durch die Aktion beeinflußte Verteilung $\rho(x,a) \in M(\mathcal{X}, \wp(\mathcal{X}))$ des Zustandes der nächsten Periode ist ein Inhalt oder ein σ-additives Wahrscheinlichkeitsmaß auf der Potenzmenge $\wp(\mathcal{X})$. $M(\mathcal{X}, \wp(\mathcal{X}))$ ist im folgenden die Menge aller Inhalte μ auf $(\mathcal{X}, \wp(\mathcal{X}))$ mit $\mu(\mathcal{X}) = 1$ **(W-Inhalt)** und $P(\mathcal{X}, \mathfrak{A})$ die Menge aller Maße ν auf $(\mathcal{X}, \mathfrak{A})$ mit $\nu(\mathcal{X}) = 1$ **(W-Maß)**.

Die obige Beschreibung führt zur folgenden Definition eines Markow-Entscheidungsmodells:

DEFINITION. $(\mathcal{X}, \mathcal{A}, \Gamma, \rho, r)$ *ist ein* **Markow-Entscheidungsmodell** **(MEM)** *mit den folgenden Eigenschaften:*

(i)	$\mathcal{X} \neq \emptyset$ *ist abzählbar*	**(Zustandsraum)**
(ii)	$\mathcal{A} \neq \emptyset$	**(Aktionenraum)**
(iii)	$\Gamma : \mathcal{X} \to \wp(\mathcal{A}) \setminus \{\emptyset\}$	**(Auswahlfunktion)**
(iv)	ρ *Übergangsinhalt von* $\mathcal{X} \times \mathcal{A}$ *nach* $\mathcal{X}$	**(Bewegungsgesetz)**
(v)	$r : \mathcal{X} \times \mathcal{A} \times \mathcal{X} \to \mathbb{R}$.	**(Einschrittauszahlungsfunktion)**

DEFINITION. *Seien* $\mathcal{Y}$ *und* $\mathcal{Z}$ *nicht leere Mengen. Ein* **Übergangsinhalt** *von* $\mathcal{Y}$ *nach* $\mathcal{Z}$ *ist eine Abbildung* $\rho : \mathcal{Y} \times \wp(\mathcal{Z}) \to [0,1]$ *mit* $\rho(y) \in M(\mathcal{Z}, \wp(\mathcal{Z}))$ $(y \in \mathcal{Y})$.

Die Auswahl der Aktionen und damit die Steuerung des Prozesses läßt sich durch eine Folge von Funktionen beschreiben, die jeder beobachteten Vergangenheit eine Aktion zuordnet. Oft wird angenommen, daß die Aktion durch ein Zufallsexperiment ausgesucht wird. Ähnlich wie in manchen statistischen Modellen hofft man, dadurch eine größere Auszahlung zu erreichen. Solche Entscheidungsfunktionen werden durch folgende Definition beschrieben:

DEFINITION.

(*i*) *Ein* **Plan** π *ist eine Folge von Abbildungen* $(\pi_i)_{i\in\mathbb{N}}$, *mit:*
π_i *ist Übergangsinhalt von* $\mathcal{X}\times(\mathcal{A}\times\mathcal{X})^{i-1}$ *nach* $\mathcal{A}$ *und*
$\pi_i(x_0,a_1..,x_i)(\Gamma(x_i))=1$.

(*ii*) *Ein Plan ist* **deterministisch**, *wenn* $\pi_i(x_0,..,x_i)$ *auf genau einen Punkt in* $\Gamma(x_i)$ *konzentriert ist.*

Ist ρ_0 ein Inhalt auf $(\mathcal{X},\mathcal{P}(\mathcal{X}))$, π ein Plan und ist ρ ein Übergangsinhalt von $\mathcal{X}\times\mathcal{A}$ nach $\mathcal{X}$, so definiert die Folge $(\rho_0,\pi_1,\rho,\pi_2,\rho...)$ einen Inhalt auf der σ-Algebra $\bigotimes_0^\infty(\mathcal{P}(\mathcal{X})\otimes\mathcal{P}(\mathcal{A}))$ von $(\mathcal{X}\times\mathcal{A})^\infty$ [cf. Dubins und Savage (1965), L.E. Dubins (1975), R. Purves und W.D. Sudderth (1973,1976)]. Dubins und Savage (1965) nennen diesen Inhalt eine **Strategie** und kürzen ihn mit σ_π ab. Er wird zur Beurteilung des Plans π dienen.

Ähnlich wie in der üblichen Integrationstheorie läßt sich mit Hilfe dieses Maßes ein Integralbegriff für Treppenfunktionen festlegen und durch gleichmäßige Konvergenz auf meßbare, beschränkte Funktionen ausdehnen. Nicht eindeutige Erweiterungen auf unbeschränkte Funktionenklassen sind möglich. Da bei den folgenden Ausführungen aber der Schwerpunkt auf eine Darstellung mit Hilfe der üblichen σ-additiven Wahrscheinlichkeitstheorie gelegt wird, verzichten wir auf eine ausführliche Diskussion der zugehörigen Integrationstheorie und verweisen auf das Buch von Dunford-Schwartz (1956), sowie die Arbeiten von R. Purves und W.D. Sudderth (1973) und H.-H. Bodewig (1979).

Sind π_i und ρ Übergangswahrscheinlichkeiten bezüglich geeigneter σ-Algebren $\mathfrak{D}$ und $\mathfrak{A}$ [d.h. $\pi_i(x_0,a_1...x_i)(\mathcal{D})$ meßbar in $(x_0,a_1...x_i)$ für alle $\mathcal{D}\subset\mathcal{A}$ und $\pi_i(x_0,a_1...x_i)(\mathcal{A})=1$; bzw. $\rho(x,a)(\mathcal{B})$ meßbar in (x,a) für alle $\mathcal{B}\subset\mathcal{X}$ $\rho(x,a)(\mathcal{X})=1$], so definiert die Folge $(\rho_0,\pi_1,\rho,\pi_2,\rho...)$ nach dem Satz von

Ionescu Tulcea ein Maß σ_π auf $\bigotimes_0^\infty(\mathcal{X}\otimes\mathcal{A})$. R. Purves und W.D. Sudderth (1976) haben gezeigt, daß unter den Meßbarkeitsvoraussetzungen dieses Maß mit dem Inhalt σ_π (eingeschränkt auf $\bigotimes_0^\infty(\mathcal{X}\otimes\mathcal{A})$) übereinstimmt. Deshalb wird hier für Inhalt und Maß das gleiche Symbol gewählt.

Aus den angeführten Gründen ist es notwendig, bei der Behandlung des $\mathcal{MEM}$ mit σ-additiven Maßen die Meßbarkeit der Pläne und der Funktion ρ zu fordern. Die Forderung der Meßbarkeit der Übergangswahrscheinlichkeiten π_i und ρ ist erfüllt, wenn π ein deterministischer Plan ist, und $\mathcal{A} = \mathcal{P}(A)$ und $\mathcal{X} = \mathcal{P}(X)$ gewählt werden. Daher kann man in die sem Fall weitere Meßbarkeitsüberlegungen außer Acht lassen. Da der Schwerpunkt dieser Arbeit auf deterministischen Plänen liegt, wird auf eine weitere Diskussion der Meßbarkeitsfragen verzichtet. Es wird sich außerdem zeigen, daß randomisierte Pläne in vielen Fällen durch deterministische Pläne ersetzt werden können. Die **Menge aller Pläne** wird mit Π bezeichnet.

Meist wird in dieser Arbeit angenommen, daß der Startpunkt x oder sogar die Vorgeschichte $(x_0, a_1, ..., a_i, x_i)$ bekannt ist. Im Fall des bekannten Anfangszustands x ist es sinnvoll, $\rho_0 = \delta_x$ zu setzen. Das von der Folge $(\delta_x, \pi_1, \rho, \pi_2, \rho...)$ induzierte Maß σ_π wird mit $\sigma_\pi[x]$ abgekürzt. Ist die Vorgeschichte $(x_0, a_0, ...a_{i-1}, x_i)$ bekannt, so bezeichnet $\pi[x_0, a_1, ..., a_i, x_i]$ den **(bedingten) Plan** π', mit:

$$\pi'_j(x'_0, a'_1, .., x'_j) := \pi_{i+j}(x_0, a_1, .., x_i, a'_1, .., x'_j) \qquad x'_0 = x_i.$$

Betrachtet man x_i als den Startpunkt eines neuen Prozesses und setzt

$$\rho'_0 := \delta_{x_i},$$

so wird durch $(\rho'_0, \pi'_1, \rho, \pi'_2, \rho...)$ eine Strategie induziert. Sie wird mit $\sigma_\pi[x_0, a_1, ..., x_i]$ bezeichnet. Es sei angemerkt, daß $\sigma_\pi[x_0, a_1, ..., x_i]$ die **bedingte Verteilung** von $\sigma_\pi[x_0]$ unter Voraussetzung $(x_0, a_1, ..., a_i, x_i)$ ist.

Zur Bewertung des Nutzens eines Plans π bis zu einem Stoppzeitpunkt t wird in den Modellen der dynamischen Optimierung $\int u_t \; d\sigma_\pi[x] \;\; [u_t(x_0 \cdots) := \sum_i^t r(x_i, a_{i+1}, x_{i+1})]$ herangezogen. Dieser Ausdruck ist das Mittel über die Summe der Einschrittauszahlungen bis zum Zeitpunkt t. Bei der Beurteilung von Prozessen entlang eines Pfades mit langer Lebensdauer wird oft der Grenzwert $\sum_i^\infty r(x_i, a_{i+1}, x_{i+1})$ genommen. Dazu muß aber gefordert werden, daß dieser Grenzwert existiert. Eine solche Konvergenzannahme ist

eine Einschränkung an den Prozeßverlauf. Prozesse, die in der Zukunft stark schwankende Auszahlungen liefern, werden nicht mehr erfaßt.

In dieser Arbeit werden Prozesse mit langer Lebensdauer betrachtet. Um auch Prozesse mit stark schwankende Auszahlungen betrachten zu können, wird als Ertragsfunktion entlang eines Pfades $\overline{\lim}_{n\in\mathbb{N}} \sum_{i=0}^{n-1} r(x_i, a_{i+1}, x_{i+1})$ herangezogen. Für den mittleren Gewinn eines Plans ergäbe sich dann

$$\int \overline{\lim_{n\in\mathbb{N}}} u_n \;\; d\sigma_\pi[x] = \int \overline{\lim_{n\in\mathbb{N}}} \sum_{i=0}^{n-1} r(x_i, a_{i+1}, x_{i+1}) \;\; d\sigma_\pi[x].$$

Angeregt durch die Arbeit von Dubins und Savage (1965) soll hier aber die folgende **Gewinnfunktion** betrachtet werden:

$$\overline{\lim_{t\in T}} \int u_t \;\; d\sigma_\pi[x] = \overline{\lim_{t\in T}} \int \sum_{i=0}^{t-1} r(x_i, a_{i+1}, x_{i+1}) \;\; d\sigma_\pi[x].$$

Dabei ist T eine noch zu spezifizierende Menge von Stoppzeiten. Diese Funktion suggeriert mehr als die erste, daß wir uns mit Systemen beschäftigen wollen, die in ferner Zukunft zu einem günstigen Zeitpunkt gestoppt werden können. Es wird sich zeigen, daß unter einer leichten Voraussetzung an den Plan π und bei geeigneter Wahl der Menge T beide Funktionen den gleichen Wert liefern.

DEFINITION.

(*i*) *Sei* $(\mathcal{A}_n)_n$ *eine aufsteigende Folge von* σ*-Algebren auf einer Menge* $\mathcal{Y}$. *Eine* **Stoppzeit** t *auf* $(\mathcal{Y}, (\mathcal{A}_n)_n)$ *ist eine Funktion* $t : \mathcal{Y}^\infty \to \mathbb{N} \cup \{\infty\}$ *mit:* $t^{-1}(n) \in \mathcal{A}_n$.

(*ii*) *Eine* **endliche Stoppzeit** t *ist eine Stoppzeit mit* $t < \infty$.

(*iii*) *Eine* **beschränkte Stoppzeit** t *ist eine Stoppzeit mit* $t < m$ *für ein* $m \in \mathbb{N}$.

Wenn nichts anderes gesagt ist, wird angenommen, daß $\mathcal{Y} = (\mathcal{X} \times \mathcal{A})^\infty$ und $\mathcal{A}_n = \bigotimes_0^n (\mathcal{X} \otimes \mathcal{A}) \times (\mathcal{X} \times \mathcal{A})^\infty$ ist. Eine Stoppzeit t läßt sich dann auf folgende Weise charakterisieren:

> Ist $(x_0, a_1, ...)$ ein Pfad und $n = t(x_0, a_1, ...)$, so gilt für jeden Pfad $(x'_0, a'_1, ...)$ mit $x_i = x'_i$ $(0 \le i \le n)$ und $a_i = a'_i$ $[1 \le i \le n,$ $t(x'_0, a'_1, ...) = n\,]$.

Im folgenden ist T_b die Menge der beschränkten Stoppzeiten, T_e die Menge der endlichen Stoppzeiten und T_∞ die Menge der $\sigma_\pi[x]$-fast-sicher endlichen Stoppzeiten ($\forall \pi \in \Pi$). Soll eine Aussage sowohl für $T = T_e$ als auch für $T = T_\infty$ gelten, so wird die Menge der Stoppzeiten mit T bezeichnet.

Eine spezielle Stoppzeit ist die **Ankunftzeit** $t_\mathcal{D}$ in einer Menge $\mathcal{D} \subset \mathcal{X}$. Sie ist die definiert durch:

$$t_\mathcal{D} := \min\{\, n \mid x_n \in \mathcal{D} \,\}.$$

$t_\mathcal{D}$ ist genau dann eine fast sicher endliche Stoppzeit im Sinne der obigen Definition, wenn $\mathcal{D}$ fast sicher erreicht wird. Ist $\mathcal{D} = \{x\}$, so wird $t_{\{x\}}$ durch t_x abgekürzt.

In dieser Arbeit wird hauptsächlich mit endlichen und fast sicher endlichen Stoppzeiten gearbeitet. Beide Mengen von Stoppzeiten bieten für die mathematische Behandlung Vorteile. Wie sich zeigen wird, ermöglichen die fast-sicher endlichen Stoppzeiten eine einfachere Darstellung der Nutzenfunktion. Demgegenüber zeigen Dubins und Savage (1965), daß sich die Menge der endlichen Stoppzeiten teilordnen läßt. Diese Eigenschaft kann man durch Anwendung der transfiniten Induktion ausnutzen. Für eine Darstellung der Teilordnung sei auf das Buch von Dubins und Savage (Paragraph 2.7) verwiesen. Ferner bieten diese Stoppzeiten den Vorteil einer größeren Verwandschaft zu den Modellen mit endlichem Horizont, was für die Anwendung interessant ist.

Zur Definition der Teilordnung führen Dubins und Savage (1965) eine Funktion α von dem Raum der endlichen Stoppzeiten in die Menge der Ordinalzahlen ein. Sie nennen $\alpha(t)$ die **Struktur** von t. Die Struktur einer Stoppzeit ist ein Maß für die Abhängigkeit der Stoppzeit von den Variablen. So haben konstante Stoppzeiten die Struktur 0, und mit einer Stoppzeit t haben alle bedingten Stoppzeiten $t[x]$ echt kleinere Struktur als t. α ist durch diese Eigenschften eindeutig bestimmt.

In dieser Arbeit werden Bedingungen für die Existenz von Plänen untersucht, deren mittlerer Gewinn möglichst groß ist. Dabei wird der **maximale mittlere Gewinn** mit

$$V(x) := \sup_{\pi \in \Pi} \overline{\lim_{t \in T}} \int u_t \quad \mathrm{d}\sigma_\pi[x]$$

abgekürzt. Das führt zur Definition von optimalen, bzw. ε-optimalen Plänen:

DEFINITION.

(*i*) *Ein Plan* π *ist* **optimal in** x, *wenn*

$$V(x) = \overline{\lim_{t \in T}} \int u_t \ \mathrm{d}\sigma_\pi[x].$$

(*ii*) *Ein Plan ist* **optimal**, *wenn er optimal in allen* x *ist.*

(*iii*) *Ein Plan ist* ε**-optimal** ($\varepsilon > 0$), *wenn*

$$V(x) < \overline{\lim_{t \in T}} \int u_t \ \mathrm{d}\sigma_\pi[x] + \varepsilon \quad \forall_x \text{ ist.}$$

Man sieht leicht, daß immer ein ε-optimaler Plan existiert, während die Existenz von optimalen Plänen nicht notwendig gesichert ist.

Für theoretische und angewandte Überlegungen ist es hilfreich zu wissen, ob ein Plan in der ganzen Zukunft optimal ist, d.h. ob jeder bedingte Plan optimal ist. Erfüllt ein Plan diese Bedingung nicht, so kann es sein, daß der beobachtete Prozeß in Zustände gelangt, in denen der Plan sehr schlechte Auszahlungen erwarten läßt. In diesem Fall wäre es dann notwendig, während des Prozeßverlaufs einen anderen Plan zu ergreifen, was in der Praxis mit Kosten verbunden ist.

DEFINITION.

(*i*) *Ein Plan* π *ist* **überall optimal in** x, *wenn jeder bedingte Plan* $\pi[x, a_1, x_1, ..., x_n]$ *optimal in* x_n *ist.*

(*ii*) *Ein Plan ist* **überall optimal**, *wenn er überall-optimal in* x *für alle* x *ist.*

(*iii*) *Ein Plan ist* **überall** ε**-optimal** ($\varepsilon > 0$), *wenn alle bedingten Pläne* $\pi[x, a_1, x_1, ..., x_n]$ *überall-*ε*-optimal in* x_n *sind.*

Ziel dieser Arbeit ist es zu untersuchen, ob optimale Pläne stationär oder wenigstens Markowsch gewählt werden können.

DEFINITION.

(*i*) *Ein Plan* π *ist ein* **Markow-Plan**, *wenn* $\pi_i(x_0, a_1, x_1.., x_i)$ *nur von* i *und dem Zustand* x_i *abhängt.*

(*ii*) *Ein Plan ist* **stationär**, *wenn* $\pi_i(x_0, ..., x_i)$ *nur von* x_i *abhängt.*

Einige Schreibweisen werden im Umgang mit Stoppzeiten benötigt. Ist eine Vorgeschichte $(x_0, ..., x_n)$ durchlaufen und ist t eine Stoppzeit mit $t(x_0, ..., x_n, h) > n$ für wenigstens ein $h \in (X \times A)^\infty$, so gilt wegen der

Definition der Stoppzeit $t(x_0, ..., x_n, h) > n$ für alle $h \in (\mathcal{X} \times \mathcal{A})^\infty$. Für eine solche Vorgeschichte $(x_0, ..., x_n)$ und ein solches t ist die Funktion

$$t[x_0, ..., x_n](h) := t(x_0, ..., x_n, h) - n$$

eine Stoppzeit. $t[x_0, ..., x_n]$ wird im folgenden als **bedingte Stoppzeit** von t unter der Bedingung $(x_0, ..., x_n)$ bezeichnet. Für jede Folge $h \in (\mathcal{X} \times \mathcal{A})^\infty$ ist $x_i(h)$ der Zustand, in dem sich das Modell zur Zeit $i \in \mathbb{N}$ befindet; ist t eine Stoppzeit, so bezeichnet $x_t(h)$ den Zustand zur Zeit $t(h)$. Analog ist $a_i(h)$ bzw. $a_t(h)$ die Aktion zur Zeit i bzw. $t(h)$. Weiter werden wir eine Vorgeschichte $(x_0, ..., x_t)$ mit p_t abkürzen.

Bei der Darstellungen von Erwartungswerten werden einige Vereinfachungen vorgenommen. Dabei wird nach folgendem Grundprinzip verfahren:

> Funktionen unter einem Integral haben, außer speziell bezeichneten, keine freien Variablen.

Der Ausdruck $\underline{\lim}_{n \in \mathbb{N}} \int u_t \; d\sigma_\pi[x]$ ist daher eine Abkürzung für $\underline{\lim}_{n \in \mathbb{N}} \int u_t(h) \; \sigma_\pi[x](dh)$ und $\int u_t + \underline{\lim}_{n \in \mathbb{N}} \int u_t \; d\sigma_\pi[x_1, ..., x_t] \; d\sigma_\pi[x]$ ist als $\int u_t(h) + \underline{\lim}_{n \in \mathbb{N}} \int u_t(h') \, \sigma_\pi[x_1(h), ..., x_t(h)](dh') \, \sigma_\pi[x](dh)$ zu interpretieren.

Für die folgenden Überlegungen ist es oft nötig, Stoppzeiten hintereinanderzuschalten. Ein solche **Hintereinanderschaltung** zweier Stoppzeiten s, t wird mit $t \oplus s$ bezeichnet. Ist $(x_0, a_1 ....) \in (\mathcal{X} \times \mathcal{A})^\infty$ ein Prozeßverlauf und $s(x_0, a_1 ...) = n$, so ist

$$(t \oplus s)(x_0, a_1 ....) := t(x_n, a_{n_1}, x_{n+1}, ...) + n.$$

Endsprechend ist $\bigoplus_i^k s_i$ die Summe von k Stoppzeiten $(s_i)_{i \leq k}$.

Oft wird im folgenden aus einer Folge von Plänen und Stoppzeiten $(\pi^n, t_n)_n$ ein Plan zusammengesetzt werden. Die Idee ist, daß man zuerst den Plan π^1 verfolgt, bis die Stoppzeit t_1 den Prozeß in einem Zustand x_{t_1} stoppt. Dann startet man den Plan π^2 ausgehend vom Zustand x_{t_1} und fährt fort bis zum Zeitpunkt $t_2 \oplus t_1$. Der **zusammengesetzte Plan** ist somit wie folgt

definiert:

$$\pi'_i(x_0\cdots x_i) := \begin{cases} \pi^1_i(x_0\cdots x_i) & \text{falls} \quad t_1(h)\geq i \\ \pi^2_{i-t_1(h)}\left(x_{t_1(h)}\cdots x_{i-t_1(h)}\right) & \text{falls} \quad (t_1\oplus t_2)(h)\geq i > t_1(h) \\ \vdots & \\ \pi^{j+1}_{i\ \ominus^j_k t_k(h)}\left(x_{\oplus^j_k t_k(h)}\cdots x_{i-\oplus^i_k t_k(h)}\right) & \\ & \text{falls} \quad \bigoplus_k^{j+1} t_k(h)\geq i > \bigoplus_k^{j} t_k(h) \end{cases}$$

In einigen Fällen, besonders im zweiten Teil dieser Arbeit, wird aus einer Folge von Markow-Plänen und Stoppzeiten $(\pi^n, t_n)_n$ ein Plan zusammengesetzt werden. Damit der zusammengesetzte Plan wieder ein Markow Plan ist, muß er etwas anders definiert werden. Unter einem **zusammengesetzten Markow-Plan** wird im folgenden ein Plan (π'_i) verstanden mit:

$$\pi'_i(x_0\cdots x_i) := \begin{cases} \pi^1_i(x_0\cdots x_i) & \text{falls} \quad t_1(h)\geq i \\ \pi^2_i(x_0\cdots x_i) & \text{falls} \quad (t_1\oplus t_2)(h)\geq i > t_1(h) \\ \vdots & \\ \pi^{j+1}_i(x_0\cdots x_i) & \text{falls} \quad \bigoplus_k^{j+1} t_k(h)\geq i > \bigoplus_k^{j} t_k(h) \end{cases}$$

2. Beziehungen zu anderen Modellen der Dynamischen Optimierung.

In der Dynamischen Optimierung werden verschiedene Modelle mit additiver Gewinnfunktion betrachtet. Neben ihnen stehen das Gambling Modell $(\mathcal{GM})\,(\mathcal{X}, \mathcal{A}, \Gamma, \rho, u)$ von Dubins und Savage (1965) und die Stoppmodelle. Das hier vorgestellte Modell verallgemeinert diese Modelle.

Die Modelle mit additiver Gewinnfunktion unterscheiden sich, wie schon erwähnt, vom $\mathcal{MCM}$ dadurch, daß die fast sichere Konvergenz der Folge (u_n) bezüglich aller zulässigen Pläne gefordert wird. Diese Forderung steht im Widerspruch zum Anliegen von Dubins und Savage, Prozesse zu betrachten, deren Erträge im Zeitablauf beliebig schwanken können.

Das Modell von Dubins und Savage (1965) betrachtet im Gegensatz zu den Modellen mit additiver Gewinnfunktion nicht die Summe von Teilerträgen, sondern mißt den Wert $u(x)$ eines Zustands x. Der Nutzen einer Folge von Zuständen $(x_0, ..., x_n)$ wird daher bei ihnen mit $u(x_n)$ angegeben und der Nutzen eines Plans mit $\overline{\lim}_{t\in T} \int u(x_t)\ \mathrm{d}\sigma_\pi[x]$. Setzt man im $\mathcal{MEM}$ $r(x,a,y) = u(y) - u(x)$, so ergibt sich $\overline{\lim}_{t\in T} \int u_t\ \mathrm{d}\sigma_\pi[x] = \overline{\lim}_{t\in T} \int u(x_t)\ \mathrm{d}\sigma_\pi[x] - u(x)$, also die Nutzenfunktion bei Dubins und Savage, weshalb das $\mathcal{GM}$ als Teilmodell des $\mathcal{MEM}$ aufgefaßt werden kann.

Einige der Modellkomponenten des $\mathcal{MEM}$ sind zur Beschreibung eines Prozesses notwendig, andere nicht. So kann man sich leicht überlegen, daß die Aufgabe von Γ durch die Funktion ρ übernommen werden könnte. Wählt man nämlich eine Funktion $f : \mathcal{X} \to \mathcal{A}$ mit $f(x) \in \Gamma(x)$ und ersetzt ρ durch:

$$\rho'(x,a) := \begin{cases} \rho(x,a) & \text{, wenn } a \in \Gamma \\ \rho(x,f(x)) & \text{, wenn } a \notin \Gamma, \end{cases}$$

so kann auf die Funktion Γ verzichtet werden. D. Blackwell (1976) zeigt, daß diese Konstruktion auch auf Modelle mit Borelschem Zustandsraum übertragbar ist.

Ferner besteht das $\mathcal{GM}$ bei Dubins und Savage nur aus dem Zustandsraum $\mathcal{X}$, der Auswahlfunktion Γ und der Auszahlungfunktion u. In dieser Arbeit wird aber aus Notationsgründen das Tupel $(\mathcal{X}, \mathcal{A}, \Gamma, \rho, u)$ als $\mathcal{GM}$ bezeichnet. Die Hinzunahme des Aktionenraums und des Übergangsgesetzes ρ in das Modell ist keine Erweiterung, sie dient nur der Anpassung an die üblichen dynamischen Optimierungsmodelle und einer leichteren Anwendung bei praktische Fragestellungen.

Einige Modelle der Dynamischen Optimierung beinhalten einen Diskontierungsfaktor β. Hier wurde er weggelassen, einmal aus Anlehnung an die Arbeit von Dubins und Savage und zum anderen da auf Untersuchungen, die einen Diskontierungsfaktor benötigen, verzichtet wurde. Ist ein Diskontierungsfaktor für Anwendungen notwendig, so kann er durch Einführen eines weiteren Zustands $*$ und Änderung des Bewegungsgesetzes ρ in

$$\rho'(x,a)(D) := \beta(x,a)\ \rho(x,a)\big(D - \{*\}\big)\ +\ (1-\beta)\ 1_*(D)$$

in das Modell integriert werden.

In einigen Modellen hängt die Einschrittauszahlungsfunktion r nur vom derzeitigen Zustand x und der Aktion a ab. Definiert man eine neue Einschrittauszahlungsfunktion $r'(x,a) := \int r(x,a,y) \quad \rho(x,a)(\ \mathrm{d}y)$, so hängt r' nur vom aktuellen Zustand und der Aktion ab. Mit Hilfe der Technik aus dem Abschnitt 'Randomisierte Pläne' läßt sich zeigen, daß die maximale mittlere Auszahlung für beide Modelle gleich ist. Wir verzichten hier auf eine Darstellung dieses Satzes.

In der Modellbeschreibung wurde erwähnt, daß die Forderung der σ-Additivität des Übergangsgesetzes ρ nicht unbedingt notwendig zur Modellformulierung ist. Der Leser wird feststellen, daß die meisten Sätze und deren Beweise auch in diesem Fall gelten. Damit alle Integrale existieren, ist es dann notwendig, daß die Funktionenfolge $(u_t)_t$ gleichmäßig beschränkt ist. Ferner ist in den Beweisen Levy's Martingal-Konvergenz-Satz durch Satz 10.3 von R. Purves und W.D. Sudderth (1973) zu ersetzen.

3. Die Nutzenfunktion.

In dieser Arbeit werden einige Bedingungen benötigt, um die Integrierbarkeit der betrachteten Funktionen zu sichern. Diese Bedigungen hängen naturgemäß davon ab, mit welcher Stoppzeitenmenge die Nutzenfunktion definiert ist. Solche Bedingungen sollen in diesem Paragraph diskutiert werden. Im wesentlichen wird die Endlichkeit aller Erwartungswerte $\int u_t \ \mathrm{d}\sigma_\pi$ gefordert. In manchen Fällen lassen sich die Probleme leichter lösen, wenn der mittlere Nutzen eines Plans π in der Form $\int \overline{\lim}_{n\in\mathbb{N}} u_n \ \mathrm{d}\sigma_\pi\,[x]$ statt durch $\overline{\lim}_{t\in T} \int u_n \ \mathrm{d}\sigma_\pi\,[x]$ dargestellt werden kann. Deshalb wird im zweiten bis vierten Teil dieses Paragraphen untersucht, unter welchen Bedigungen beide Ausdrücke gleich sind. Dabei werden Sätze von R. Chen (1976), Engelbert-Engelbert (1979) und Purves-Sudderth (1973, 1976) benutzt. Diese Sätze sind im Anhang dargestellt und sind dort, soweit nötig, bewiesen.

Es wird gezeigt, daß bei der Betrachtung des Falls $T = T_\infty$ keine zusätzliche Voraussetzung gemacht werden muß. Im Falle $T = T_e$ muß dagegen noch die gleichförmige Integrierbarkeit von $(u_t)_t$ gefordert werden. Im vierten Teil werden dann Anmerkungen zu MCM-Modellen mit endlich additiven Maßen gemacht.

Die folgenden Bedingungen werden generell vorausgesetzt:

(A1) $$\forall_x \forall_\pi \forall_{t\in T} \quad \int u_t^+ \ \mathrm{d}\sigma_\pi[x] < \infty^*$$

(A2) $$\forall_x \quad \sup_\pi \overline{\lim_{t\in T}} \int u_t^+ \ \mathrm{d}\sigma_\pi[x] < \infty$$

(A1) ist stärker ist als:

(A1') $$\forall_x \forall_\pi \forall_{t\in T_b} \quad \int u_t^+ \ \mathrm{d}\sigma_\pi[x] < \infty$$

und aus Satz 15.3(i) folgt, daß im Falle $T = T_b$ (A2) stärker ist als:

(A2') $$\forall_x \quad \sup_\pi \overline{\lim_{t\in T_\infty}} \int u_t^+ \ \mathrm{d}\sigma_\pi[x] < \infty$$

BEMERKUNG. *Für $T_b \subseteq T \subseteq T_\infty$ und ein $t \in T_b$ gilt:*

(i) $$\sup_{t\le s, s\in T_b} \int u_s^+ \ \mathrm{d}\sigma_\pi[x] = \sup_{t\le s, s\in T} \int u_s^+ \ \mathrm{d}\sigma_\pi[x] = \sup_{t\le s, s\in T_\infty} \int u_s^+ \ \mathrm{d}\sigma_\pi[x]$$

(ii) $$\overline{\lim_{t\in T_b}} \int u_t^+ \ \mathrm{d}\sigma_\pi[x] \ge \overline{\lim_{t\in T}} \int u_t^+ \ \mathrm{d}\sigma_\pi[x] \ge \overline{\lim_{t\in T_\infty}} \int u_t^+ \ \mathrm{d}\sigma_\pi[x]$$

BEWEIS: (i) Die Ungleichung '$\le$' ist klar. Ist $t \in T_\infty$, so gilt nach dem Lemma von Fatou

$$\int u_t^+ \ \mathrm{d}\sigma_\pi[x] \le \underline{\lim_{n\in N}} \int u_{t\wedge n}^+ \ \mathrm{d}\sigma_\pi[x]$$

$$\le \sup_{t\in T_b} \underline{\lim_{n\in N}} \int u_t^+ \ \mathrm{d}\sigma_\pi[x].$$

(ii) folgt direkt aus (i). ∎

Aus den obigen Bedingungen folgt nun die Integrierbarkeit der benötigten Funktionen:

---

* $g^+ := \max\{g, 0\}$, $g^- := \max\{-g, 0\}$

3.1 LEMMA. *Sei $T_b \subseteq T \subseteq T_\infty$, und gelte (A1) und (A2), dann gilt:*

(*i*) $V < \infty$

(*ii*) $u_s + \overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_\pi[p_s]$ *ist $\sigma_\pi[x]$-integrierbar für alle $s \in T_\infty$, $\pi \in \Pi$*

(*iii*) *Ist $\overline{\lim}_{t\in T} \int u_t \ \mathrm{d}\sigma_\pi[x] > -\infty$ und ist $g : \mathbb{R} \to \mathbb{R}$ eine bezüglich aller Bildmaße von u_s auf $\mathbb{R}$ integrierbare Funktion, so gilt:*

$$\overline{\lim_{t\in T}} \int g(u_s p_s + u_t) \ \mathrm{d}\sigma_\pi[p_s]^{**} \quad \text{ist } \sigma_\pi[x]\text{-integrierbar}$$

für alle $s \in T_\infty$, $\pi \in \Pi$

(*iv*) *$u_s + V(x_s)$ ist $\sigma_\pi[x]$-quasi-integrierbar für alle $s \in T_\infty$, $\pi \in \Pi$*

(*v*)
$$\overline{\lim_{t\in T}} \int_{\mathcal{D}} u_t \ \mathrm{d}\sigma_\pi[x] = \int_{\mathcal{D}} u_s + \overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_\pi[p_s] \ \mathrm{d}\sigma_\pi[x]$$

für alle $s \in T$ und alle $\mathcal{D}$ aus der von $(x_{t\wedge n})_n$ aufgespannten σ-Algebra auf $(\mathcal{X} \times \mathcal{A})^\infty$

(*vi*) *Ist $\overline{\lim}_{t\in T} \int u_t \ \mathrm{d}\sigma_\pi[x] > -\infty$ und ist $g : \mathbb{R} \to \mathbb{R}$ eine bezüglich aller Bildmaße von u_s auf $\mathbb{R}$ integrierbare Funktion, so gilt:*

$$\overline{\lim_{t\in T}} \int_{\mathcal{D}} g(u_s p_s + u_t) \ \mathrm{d}\sigma_\pi[x] = \int \overline{\lim_{t\in T}} \int_{\mathcal{D}} g(u_s p_s + u_t) \ \mathrm{d}\sigma_\pi[p_s] \ \mathrm{d}\sigma_\pi[x]$$

für alle $s \in T$ und alle $\mathcal{D}$ aus der von $(x_{t\wedge n})_n$ aufgespannten σ-Algebra auf $(\mathcal{X} \times \mathcal{A})^\infty$

(*vii*) $u_s + \overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_\pi[p_s]$ *ist ein Martingal in T.*

BEWEIS: (*i*) folgt aus A2.

(*ii*) Aus der Quasiintegrabiltität von u_τ für alle $\tau \in T$ folgt die Integrierbarkeit von $u_s + \int u_t \ \mathrm{d}\sigma_\pi[p_s]$ für alle $t \in T$. Aus der Existenz eines $t \in T$ mit:

$$u_s + \overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_\pi[p_s] - \varepsilon \leq u_s + \int u_t \ \mathrm{d}\sigma_\pi[p_s]$$

folgt dann (*ii*).

** *Ist x ein Zustand, $g : (\mathcal{X} \times \mathcal{A})^\infty \to \mathbb{R}$, so ist $gx(a'_0, \ldots) := g(x, a'_0, \ldots)$. Analog ist gp für $p \in (\mathcal{X} \times \mathcal{A})^n$ definiert.*

(*iii*) Aus $\overline{\lim}_{t\in T}\int u_t \;\mathrm{d}\sigma_\pi[x] > -\infty$ folgt die Integrierbarkeit von u_τ für alle τ, die größer als ein τ' sind. Wegen der Integrierbarkeit von g sind daher alle Funktionen $g\circ u_t$*** integrierbar. Daher ist der bedingte Erwartungswert $\int g\circ u_\tau p_s \;\mathrm{d}\sigma_\pi[p_s]$ für alle $\tau > s$ integrierbar. Die Integrierbarkeit des $\overline{\lim}$ folgt nun aus der Existenz zweier Stoppzeiten τ, t mit

$$\int g\circ u_t p_s \;\mathrm{d}\sigma_\pi[p_s] - \varepsilon \le \overline{\lim_{t\in T}}\int g\circ u_t p_s \;\mathrm{d}\sigma_\pi[p_s]$$
$$\le \int g\circ u_\tau p_s \;\mathrm{d}\sigma_\pi[p_s] + \varepsilon.$$

(*iv*) In jedem Punkt x existiert ein Plan π_x mit:

$$V(x) \le \int u_t \;\mathrm{d}\sigma_{\pi_x}[x] + \varepsilon.$$

Ist π ein vorgegebener Plan und π' der Plan, der bis zum Zeitpunkt s π verfolgt und dann π_{x_s} ergreift, dann gilt für π'

$$-\infty < \int u_s + \overline{\lim_{t\in T}}\int u_t \;\mathrm{d}\sigma_\pi[p_s] \;\mathrm{d}\sigma_\pi[x] \qquad \le \int u_s + V(x_s) \;\mathrm{d}\sigma_\pi[x]$$

$$\le \int u_s + \overline{\lim_{t\in T}}\int u_t \;\mathrm{d}\sigma_\pi[p_s] \;\mathrm{d}\sigma_\pi[x] + \varepsilon < \infty,$$

womit die Quasiintegrierbarkeit der Funktion aus (*ii*) folgt.

Teil (*v*) folgt aus (*iv*) und die Teile (*vi*) und (*vii*) aus den Regeln für den $\overline{\lim}$. ∎

Ist für alle $s \in T$ und $\pi \in \Pi$

(A3) $$\int \left[\overline{\lim_{t\in T}}\int u_t \;\mathrm{d}\sigma_\pi[p_s]\right]^+ \mathrm{d}\sigma_\pi[x] < \infty,$$

erfüllt, dann gilt (wegen Bedingung A1 und Definition von V)

(*viii*) $$\int V(x_s) \;\mathrm{d}\sigma_\pi[x] < \infty \quad \text{für alle } s \in T, \pi \in \Pi.$$

Die Bedingung

(A3') $$\forall_\pi \forall_s \in T \quad \int u_s \;\mathrm{d}\sigma_\pi[x] > -\infty$$

impliziert A3. Auf der anderen Seite gilt für einen Plan π mit den Eigenschaften

*** $\circ$ bezeichnet die Hintereianderschaltung von Abbildungen

(A3") $$\overline{\lim_{t\in T}}\int u_t \;\; d\sigma_\pi[x] > -\infty \quad \text{und A3}$$

die Bedingung A3'.

Die Bedingung A3 ist keine wesentliche Einschränkung. Man kann nämlich ohne Beschränkung annehmen, daß $V > -\infty$ für alle x gilt. Aus dieser Eigenschaft folgt, daß in jedem Startpunkt ein Plan π existiert mit: $\overline{\lim}_{t\in T}\int u_t \;\; d\sigma_\pi[x] > -\infty$. Für diesen Plan und für jedes große $t \in T$ gilt dann $\int u_t \;\; d\sigma_\pi[x] > -\infty$. Die Bedingung A3', die A3 impliziert, ist daher etwas stärker als $V > -\infty$.

Um die Integrabilität der benötigten Funktionen zu sichern, wird im folgenden grundsätzlich **A1 und A2** vorausgesetzt. Meist wird zusätzlich $\overline{\lim}_{t\in T}\int u_t \;\; d\sigma_\pi[x] > -\infty$ gefordert.

3a. Der Fall fast sicher endlicher Stoppzeiten. Für viele Untersuchungen ist es sehr hilfreich, wenn gilt:

$$\overline{\lim_{t\in T_\infty}}\int u_t \;\; d\sigma_\pi[x] = \int \overline{\lim_{n\in N}} u_n \;\; d\sigma_\pi[x].$$

Die Richtigkeit dieser Gleichung soll nun zuerst für den Fall $T = T_\infty$ untersucht werden. Wie aus einer Arbeit von Engelbert-Engelbert (1979, s. Paragraph 15a.) hervorgeht, gilt die Gleichung unter den Voraussetzungen:

(E1) $$\overline{\lim_{n\in N}} u_n \qquad \text{ist } \sigma_\pi[x]\text{-quasi-integrierbar } \forall_{\pi,x}$$

(E2) $$\overline{\lim_{t\in T_\infty}}\int u_t \;\; d\sigma_\pi[x] < \infty \;\forall_{\pi,x}$$

Diese Voraussetzungen sind eine Verallgemeinerung von W.D. Sudderth (1971b) und R. Chen (1976). In beiden Arbeiten wird mindestens die gleichmäßige Integrierbarkeit von (u_t) gefordert.

Aus den Überlegungen von Engelbert-Engelbert folgt:

3.2 LEMMA. *Ist* $\overline{\lim}_{n\in N} u_n$ $\sigma_\pi[x]$*-quasi-integrierbar, dann gilt:*

$$\overline{\lim_{t\in T_\infty}}\int u_t \;\; d\sigma_\pi[x] = \int \overline{\lim_{n\in N}} u_n \;\; d\sigma_\pi[x].$$

Soll die Vertauschbarkeit von $\overline{\lim}_{t\in T_\infty}$ und Integral erreicht werden, so wird im folgenden angenommen, daß **E1** zusätzlich zu den Generalveraussetzungen A1 und A2 erfüllt ist.

3b. Der Fall endlicher Stoppzeiten. Im Fall beschränkter Stoppzeiten, $T = T_b$, implizieren die Bedingungen (cf. R. Chen (1976), s. Paragraph 15b):

(C1) $(u_t^+)_{t\in T_b}$ ist eine gleichförmig $\sigma_\pi[x]$-integrierbare Menge von Funktionen $\forall_{\pi,x}$

und

(C2) $(u_n^-)_{n\in\mathbb{N}}$ ist eine gleichförmig $\sigma_\pi[x]$-integrierbare Menge von Funktionen $\forall_{\pi,x}$,

daß $\overline{\lim}_{t\in T_b}$ und Integral vertauschbar sind (d. h. $\overline{\lim}_{t\in T_b}\int u_t \ \mathrm{d}\sigma_\pi[x] = \int \overline{\lim}_{n\in\mathbb{N}} u_n \ \mathrm{d}\sigma_\pi[x]$). Diese Folgerung ist sogar für beliebige Mengen von Stoppzeiten T ($T_b \subset T \subset T_\infty$) richtig, wenn für für alle $t \in T$, $t_i \in T_\infty$ und $t_b \in T_b$ $\quad t \wedge t_i$ und $t \vee t_b$ in T liegen.

Insbesondere gilt, daß $\overline{\lim}_{n\in\mathbb{N}} u_n$ integrierbar ist (cf. Bedingung E1). Weiter zeigt Lemma 15.4, daß Bedingung C1 äquivalent zu C1' ist:
(C1') $(u_t^+)_{t\in T_\infty}$ ist eine gleichförmig integrierbare Menge von Funktionen.
Die Bedigung C1 impliziert A1 und E2.

Aus Satz 15.2 und dem oben Gesagten folgt unter den Bedingungen C1 und C2, daß gilt:

$$\overline{\lim_{t\in T_e}}\int u_t \ \mathrm{d}\sigma_\pi[x] = \overline{\lim_{t\in T_\infty}}\int u_t \ \mathrm{d}\sigma_\pi[x] = \int \overline{\lim_{n\in\mathbb{N}}} u_n \ \mathrm{d}\sigma_\pi[x] < \infty.$$

Das Lemma von Fatou impliziert nun:

$$\int \overline{\lim_{n\in\mathbb{N}}} - u_n^- \ \mathrm{d}\sigma_\pi[x] \geq \overline{\lim_{n\in\mathbb{N}}}\int - u_n^- \ \mathrm{d}\sigma_\pi[x] > -\infty,$$

woraus mit R. Chen (1976, Theorem 2.1(3)) die Endlichkeit von $\int \overline{\lim}_{n\in\mathbb{N}} u_n \ \mathrm{d}\sigma_\pi[x]$ folgt.

Will man im Fall endlicher Stoppzeiten die Vertauschbarkeit von $\overline{\lim}_{t\in T_b}$ und Integral sicherstellen, ist es deshalb nützlich, die Bedingungen **C1 und C2** vorauszusetzen.

3c. Der Fall endlich additiver Maße. Unter der Voraussetzung, daß $(u_t)_{t\in T_e}$ $\sigma_\pi[x]$-fast-sicher gleichmäßig beschränkt ist, lassen sich die Integraldarstellungen im ersten Teil des Abschnitts ebenfalls für endlich additive Maße beweisen. Ferner zeigen R. Purves und W.D. Sudderth (1973, Theorem 10.4), daß in diesem Fall die Vertauschbarkeit von Limes und Integral gilt. Als Menge der Stoppzeiten ist dann T_e zu wählen.

4. Randomisierte Pläne.

Das hier vorgestellte Modell erlaubt die Behandlung randomisierter Entscheidungen. Wie sich in der Statistik zeigt, kann der Einsatz solcher Entscheidungsfunktionen sinnvoll sein. Auf der anderen Seite konnte in dynamischen Optimierungsmodellen bisher gezeigt werden, daß durch deterministische Pläne gleich gute Gewinne erzielt werden können wie durch randomisierte Pläne (cf. D. Blackwell (1964)). Es wird gezeigt werden, daß zu einem gegebenen randomisierten Plan ein gleich guter deterministischer Plan existiert.

Darüber hinaus soll untersucht werden, ob man durch sogenannte Markow-Mischung die Auszahlung verbessern kann. Durch Markow-Mischung eines deterministischen Plans erhält man einen randomisierten Markow-Plan. Sie ist dadurch charakterisiert, daß sowohl der ursprüngliche Plan wie auch seine Markow-Mischung gleiche Randverteilungen auf die Aktionen- und Zustandsräume haben. In vielen Modellen der dynamsichen Optimierung kann gezeigt werden, daß ein Plan die gleiche mittlere Auszahlung hat wie seine Markow-Mischung. Durch D. Kadelka und Th.P. Hill wurde die Frage aufgeworfen, ob das auch für das Modell von Dubins und Savage und das $\mathcal{MEM}$ gilt. An Hand eines Beispiels wird gezeigt, daß solche Mischungen wesentlich geringere Auszahlung liefern können als der ursprüngliche deterministische Plan.

4.1 LEMMA. *Ist $t \in T_e$, $g : \mathcal{X} \to \mathbb{R}$ eine für alle $a \in \Gamma$, $x \in \mathcal{X}$ $\rho(x,a)$-integrierbare Funktion und ist $\varepsilon > 0$, dann existiert zu jedem randomisierten Plan π ein deterministischer Plan δ mit*

$$\int u_t + g(x_t) \;\; \mathrm{d}\sigma_\pi[x] \leq \int u_t + g(x_t) \;\; \mathrm{d}\sigma_\delta[x] + \varepsilon.$$

BEWEIS: Der Beweis benutzt transfinite Induktion über die Strukur von $t \in T_e$. Ist $t \in \mathbb{N}$, dann gilt die Behauptung für alle ε (Induktion).
Induktionsschluß für die transfinite Induktion: Nimmt man an, daß für ein $t \in T_e \setminus \mathbb{N}$, für alle ε und alle $s \in T_e$ mit Struktur von s kleiner als Struktur von t ein deterministischer Plan δ_s existiert, mit

$$\int u_s + g(x_s) \;\; \mathrm{d}\sigma_\pi[x] \leq \int u_s + g(x_s) \;\; \mathrm{d}\sigma_{\delta_s}[x] + \varepsilon,$$

dann folgt die Behauptung für t direkt aus:

$$\int u_t + g(x_t) \;\; d\sigma_\pi[x]$$

$$= \int \quad r(x, a_1, x_1) + \int u_{t[x]} + g(x_{t[x]}) \;\; d\sigma_\pi[x, x_1] \;\; d\sigma_\pi[x],$$

da die Struktur von $t[x]$ echt kleiner als die von t ist. ∎

4.2 SATZ. *Ist δ ein randomisierter Plan und sind C1 und C2 erfüllt, dann existiert für ein ε ein deterministischer Plan π mit*

$$\overline{\lim_{t\in T_e}} \int u_t \;\; d\sigma_\delta[x] \leq \overline{\lim_{t\in T_e}} \int u_t \;\; d\sigma_\pi[x] + \varepsilon.$$

BEWEIS: Die Idee des Beweises ist, mit Hilfe des obigen Lemmas einen determistischen Plan zu konstruieren, der stückweise besser als der Plan δ ist.

Sei $\quad t_\varepsilon := \min\left\{ n \geq 1 \;\middle|\; \overline{\lim_{t\in T_e}} \int u_t \;\; d\sigma_\delta[x_0 ... x_n] < \varepsilon \right\}.$

Wegen der Vertauschbarkeit von $\overline{\lim}$ und Integral und wegen Levy's Martingalsatz konvergiert der bedingte Erwartungswert $u_n + \overline{\lim}_{t\in T_e} \int u_t \;\; d\sigma_\delta[x_0, ..., x_n]$ gegen $\lim_{n\in\mathbb{N}} u_n$, weshalb $\sigma_\delta[x](t_\varepsilon < \infty) = 1$ folgt.

Wegen $\overline{\lim}_{t\in T_e} \int u_t \;\; d\sigma_\delta[x_0 ... x_{t_\varepsilon\wedge n}] = u_{t_\varepsilon\wedge n} + \overline{\lim}_{t\in T_e} \int u_t \;\; d\sigma_\delta[x_0 ... x_{t_\varepsilon\wedge n}] - u_{t_\varepsilon\wedge n}$ gilt:

$$\left(\overline{\lim_{t\in T_e}} \int u_t \;\; d\sigma_\delta[x_0 ... x_{t_\varepsilon\wedge n}] \right)_n \quad \text{ist gleichförmig integrierbar.}$$

Aus der Definition von t_ε folgt:

$$\left[\overline{\lim_{t\in T_e}} \int u_t \;\; d\sigma_\delta[x_0 ... x_{t_\varepsilon\wedge n}] \geq \varepsilon \right] = \left[t_\varepsilon > n \right],$$

weshalb nach Satz 15.3(*ii*) gilt:

$$0 \leq \underline{\lim_{s\in T_b}} \int 1_{[t_\varepsilon > s]} \overline{\lim_{t\in T_e}} \int u_t \;\; d\sigma_\delta[x_0 ... x_{t_\varepsilon\wedge s}] \;\; d\sigma_\delta[x]$$

$$\leq \int \underline{\lim_{n\in\mathbb{N}}} 1_{[t_\varepsilon > n]} \overline{\lim_{t\in T_e}} \int u_t \;\; d\sigma_\delta[x_0 ... x_{t_\varepsilon\wedge n}] \;\; d\sigma_\delta[x]$$

$$= 0.$$

Sei $s \in T_b$ mit:

$$\int 1_{[t_\varepsilon > s]} \overline{\lim_{t\in T_e}} \int u_t \;\; d\sigma_\delta[x_0 ... x_{t_\varepsilon\wedge s}] \;\; d\sigma_\delta[x] \leq \varepsilon,$$

dann existiert ein deterministischer Plan π mit:

$$\overline{\lim_{t\in T_\varepsilon}}\int u_t \;\; \mathrm{d}\sigma_\delta\,[x]$$

$$\leq \int \; u_{t_\varepsilon\wedge s} - \left\{1_{[t_\varepsilon\leq s]}\overline{\lim_{t\in T_\varepsilon}}\int u_t \;\; \mathrm{d}\sigma_\delta\,[x_0...x_{t_\varepsilon\wedge s}]\right\}^- \;\; \mathrm{d}\sigma_\pi\,[x] + 3\varepsilon.$$

Sei δ und ε wie in der Voraussetzung des Lemmas: $\{\varepsilon_n\}$ eine Folge von positiven reellen Zahlen mit $\sum\varepsilon_n \leq \varepsilon$. Für den Startpunkt x existiert nach Lemma 4.1 ein determinitischer Plan π_1 und eine endliche Stoppzeit s_1 mit:

$$\overline{\lim_{t\in T_\varepsilon}}\int u_t \;\; \mathrm{d}\sigma_\delta\,[x]$$

$$\leq \int u_{s_1\wedge t_{\varepsilon_1}} - \left\{1_{[t_\varepsilon\leq s_1]}\overline{\lim_{t\in T_\varepsilon}}\int u_t \;\; \mathrm{d}\sigma_\delta\,[x_0...x_{s_1\wedge t_\varepsilon}]\right\}^- \;\; \mathrm{d}\sigma_\pi\,[x] + 3\varepsilon_1.$$

Induktiv läßt sich daher eine Folge stationärer Pläne π_i und Stoppzeiten s_i konstruieren mit $(p_n := (x_0...x_n))$:

$$\overline{\lim_{t\in T_\varepsilon}}\int u_t \;\; \mathrm{d}\sigma_\delta\,[p_{s_n^*}] \;\leq$$

$$\int u_{s_{n+1}\wedge t_{\varepsilon_{n+1}}} - \left\{1_{[t_\varepsilon\leq s_n+1]}\overline{\lim_{t\in T_\varepsilon}}\int u_t \;\; \mathrm{d}\sigma_\delta\left[p_{s^*_{n+1}}\right]\right\}^- \;\; \mathrm{d}\sigma_{\pi_{n+1}}\,[p_{s_n^*}] + 3\varepsilon_{n+1},$$

wobei $s_n^* := \bigoplus_j^n s_{j\wedge t_{\varepsilon_j}+s^*_{j-1}}$ $(s_0^* := 0)$.

Sei π der zusammengesetzte Plan von (π_n, s_n) (s.Paragraph 1). Durch Induktion folgt dann für alle $n \in \mathbb{N}$:

$$\overline{\lim_{t\in T_\varepsilon}}\int u_t \;\; \mathrm{d}\sigma_\delta\,[x]$$

$$\leq \int u_{s_n^*} - \left\{1_{[t_{\varepsilon_n}\leq s_n^*]}\overline{\lim_{t\in T_\varepsilon}}\int u_t \;\; \mathrm{d}\sigma_\delta\left[p_{s_n^*}\right]\right\} \;\; \mathrm{d}\sigma_\pi\,[x] + 3\sum_{i=1}^n \varepsilon_j$$

Der Beweis erfolgt durch Induktion. Der Induktionschritt lautet wie folgt:

$$\overline{\lim_{t\in T_\varepsilon}}\int u_t \;\; \mathrm{d}\sigma_\delta\,[x]$$

$$\leq \int u_{s_n^*} - \left\{1_{[t_\varepsilon\leq s_n^*]}\overline{\lim_{t\in T_\varepsilon}}\int u_t \;\; \mathrm{d}\sigma_\delta\,[p_{s_n^*}]\right\}^- \;\; \mathrm{d}\sigma_\pi\,[x] + 3\sum_{i=1}^n \varepsilon_j$$

$$\leq \int u_{s_n^*} - \left\{ \overline{\lim_{t \in T_e}} \int u_t \ \mathrm{d}\sigma_\delta \left[p_{s_n^*} \right] \right\}^- \ \mathrm{d}\sigma_\pi [x] + 3 \sum_{i=1}^{n} \varepsilon_j$$

$$\leq \int u_{s_n^*} + \overline{\lim_{t \in T_e}} \int u_t \ \mathrm{d}\sigma_\delta \left[p_{s_n^*} \right] \ \mathrm{d}\sigma_\pi [x] + 3 \sum_{i=1}^{n} \varepsilon_j$$

$$\leq \int u_{s_n^*} + \int u_{s_{n+1} \wedge t_e} - \left\{ \overline{\lim_{t \in T_e}} \int u_t \ \mathrm{d}\sigma_\delta \left[p_{s_{n+1}^*} \right] \right\}^- \ \mathrm{d}\sigma_\pi [p_{s_n}] \ \mathrm{d}\sigma_\pi [x]$$
$$+3 \sum_{i=1}^{n} \varepsilon_j$$

$$\leq \int u_{s_{n+1}^*} - \left\{ 1_{[t_e \leq s_{n+1}^*]} \overline{\lim_{t \in T_e}} \int u_t \ \mathrm{d}\sigma_\delta \left[p_{s_{n+1}^*} \right] \right\}^- \ \mathrm{d}\sigma_\pi [x] + 3 \sum_{i=1}^{n+1} \varepsilon_j,$$

wobei der letzte Schritt aus der Konstruktion des Plans π folgt.

Aus dem vorhergehenden folgt:

$$\overline{\lim_{t \in T_e}} \int u_t \ \mathrm{d}\sigma_\delta [x] \leq \int u_{s_n^*} \ \mathrm{d}\sigma_\pi [x] + 3 \sum_{i=1}^{n} \varepsilon_j$$

$$\leq \sup_{t \geq n} \int u_t \ \mathrm{d}\sigma_\pi [x] + 3 \sum_{i=1}^{n} \varepsilon_j$$

und durch Grenzübergang in n folgt dann:

$$\overline{\lim_{t \in T_e}} \int u_t \ \mathrm{d}\sigma_\delta [x] \leq \int u_t \ \mathrm{d}\sigma_\pi [x] + \varepsilon,$$

was zu beweisen war. ∎

Der obige Satz und Beweis ist für endlich additive $\mathcal{M}$ richtig. Dabei muß man die Vorrausetzung machen, daß $(u_t)_t$ gleichmäßig beschränkt ist. Beim Beweis muß Levy's Martingalsatz durch den Martingalsatz von R. Purves und W.D. Sudderth (1973, Theorem 10.3, s. Anhang ersetzt werden.

Wir wollen uns nun der Frage zuwenden, ob deterministische Pläne durch randomisierte Pläne mit einfacherer Struktur ohne Ertragsverlust ersetzt werden können. Insbesondere stellt sich die Frage, ob die Markow-Mischung eines Plans genauso gut ist wie der ursprüngliche Plan. Das folgende Beispiel zeigt, daß diese Vermutung nicht zutrifft. In diesem Beispiel wird ein Plan angegeben, mit dem ein Ziel g unendlich oft erreicht wird. Seine Markow-Mischung

erreicht das Ziel aber mit Wahrscheinlichkeit 1 nur endlich oft. Das Beispiel ist eine Erweiterung eines Beispiels von Th.P. Hill.

DEFINITION. *Ist π ein randomisierter Plan, so ist der Plan π' mit ($a \in \mathcal{A}$):*

$$\pi'_i(x_0, a_1 \cdots x_i)(\{a\}) = \frac{\int \pi(x_0, a'_1, x'_1 \cdots x'_{i-1}, a'_i, x_i)(\{a\}) \quad \sigma_\pi[x_0]\left(d(a'_1, x'_1, a'_2 \cdots)\right)}{\sigma_\pi[x_0]\left(\ x_i \text{ ist } i\text{-ter Zustand }\right)}$$

die **Markow-Mischung** *von π.*

$\pi'_i(x_0, a_1 \cdots x_i)(\{a\})$ ist die bedingte Wahrscheinlichkeit, eine Aktion a zu ergreifen unter der Bedingung x_i, sie ist unabhängig von $(a_1, ..., x_{i-1}, a_i)$.

BEISPIEL. *Der gesuchte Plan wird aus einer Folge von Bäumen $(\Psi_n)_{n \in \mathbb{N}}$ zusammengesetzt. Den ersten Baum erhält man aus dem in Bild 1 dargestellten (endlichen) Plan.*

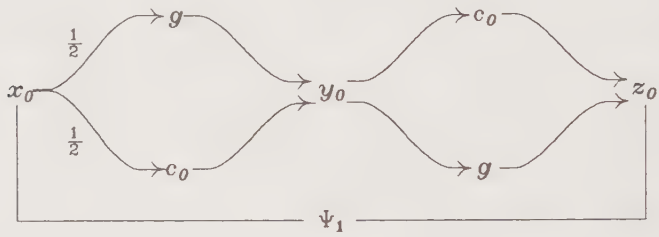

Bild 1

In diesem Plan geht das System von x_0 mit Wahrscheinlichkeit 1/2 nach g und mit Wahrscheinlichkeit 1/2 nach c_0 und dann nach y_0. Wurde im ersten Schritt g berührt, so bewegt sich das System in Schritt 4 nach c_0, sonst nach g. In Schritt 4 erreicht das System z_0. Von z_0 kann das System in jeden Zustand x_n gelangen.

Der obige Plan erreicht g mit Wahrscheinlichkeit 1, während die Markow-Mischung g mit Wahrscheinlichkeit 3/4 erreicht. Dieser Baum wird im folgenden Ψ_1 genannt, und die Wahrscheinlichkeit, g mit der Markow-Mischung von Ψ_1 zu erreichen, wird mit p_1 ($p_1 = 3/4$) abgekürzt. Die Länge von Ψ_1 wird mit $\ell(\Psi_n)$ bezeichnet, dabei gilt $\ell(\Psi_1) = 4$.

Angenommen, die Bäume $\{\Psi_1, ..., \Psi_n\}$ sind konstruiert, dann erhält man den nächsten Baum Ψ_{n+1} wie in Bild 2 dargestellt.

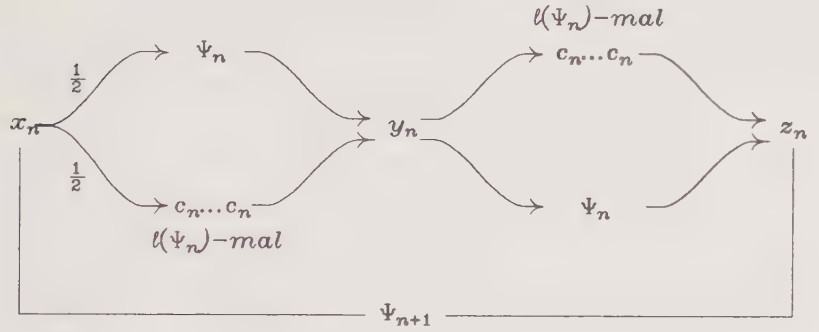

Bild 2

In diesem Plan erreicht man von a_n mit Wahrscheinlichkeit 1/2 den Baum Ψ_n und mit Wahrscheinlichkeit 1/2 $\ell(\Psi_n)$-mal den Zustand c_n. Von y_n erreicht man den Baum Ψ_n, wenn man zu Beginn den Zustand c_n erreicht hat, anderenfalls geht man $\ell(\Psi_n)$-mal durch den Zustand c_n. Von z_n kann man jeden Zustand x_i erreichen.

Für die Wahrscheinlichkeit p_{n+1} gilt:

$$p_{n+1} = 0.5\, p_n + 0.5\,(1 - 0.5\, p_n) p_n = p_n - 0.25\, p_n^2.$$

Wie die folgende Überlegung zeigt, konvergiert p_n gegen 0:
p_n ist streng monoton fallend. Angenommen, p_n konvergiert gegen eine Konstante $\alpha \neq 0$,. dann gilt $\alpha = \alpha - 0.25\,\alpha^2$. Das ist aber nur für $\alpha = 0$ richtig.

Der folgende Plan erreicht g mit Wahrscheinlichkeit 1 unendlich oft, seine Markow-Mischung erreicht g aber höchstens mit Wahrscheinlichkeit 0.5 unendlich oft.
Sei ε_i mit $\sum_{i=1}^{\infty} \varepsilon_i < 0.5$ und $(\Psi_{n_i})_{i \in \mathbb{N}}$ eine Folge von Bäumen mit $p_{n_i} \leq \varepsilon_i$. Der zusammengesetzte Plan $\{\Psi_{n_1}, \Psi_{n_2}, ...\}$ erreicht g mit Wahrscheinlichkeit 1 unendlich oft, aber die zugehörige Markow-Mischung erreicht g höchstens mit Wahrscheinlichkeit 0.5.

Es wurde gezeigt, daß unter bestimmten Vorraussetzungen ein randomisierter Plan durch einen deterministischen ersetzt werden kann. Deshalb wird im folgenden nur noch mit deterministischen Plänen gearbeitet. Π sei daher von nun an die Menge der **deterministischen Pläne**.

II Eigenschaften optimaler Pläne

Dubins und Savage schlugen vor, die Suche nach optimalen Plänen in zwei Teile zu zerlegen. Zum Verständnis ihrer Idee kann man folgendes Gedankenmodell zur Hilfe nehmen: Angenommen, man kennt einen guten Plan, den man ab einem gewissen Zeitpunkt benutzen darf, so reicht es, bis zu diesem Zeitpunkt einen Plan zu verfolgen, der möglichst wenig verschenkt, also werterhaltend ist. Diese Forderung kann man für jeden beliebigen Zeitpunkt aufstellen. Auf der anderen Seite muß ein guter Plan in ferner Zukunft optimal sein. Dubins und Savage (1965) nennen einen Plan, der die erste Bedingung erfüllt, 'thrifty' (werterhaltend), die zweite Bedingung 'equalizing' (egalisierend). In diesem Kapitel sollen die Eigenschaften werterhaltend und egalisierend charakterisiert werden. Es sollen äquivalente Eigenschaften angegeben werden.

5. Optimalitätsbedingungen.

Wählt man zuerst eine beliebige Aktion und steuert das System dann optimal, so ist das Ergebnis sicherlich selten optimal. Diese Eigenschaft wird in der Literatur meist als 'Exzessivität' der Funktion V bezeichnet. Aus ihr leitet sich die Supermartingaleigenschaft der Funktionenfolge $(u_t + V(x_t))_t$ ab.

Definition. *Eine Funktion* $f : \mathcal{X} \to \mathbb{R} \cup \{-\infty\}$ *ist* **exzessiv** *im* $\mathcal{M}_{\mathcal{C}}\mathcal{M}$, *wenn für alle* $x \in \mathcal{X}$, $a \in \Gamma$

$$\int f^+(y) \quad \rho(x,a)(dy) < \infty \qquad \textit{und}$$

$$f(x) \geq \int r(x,a,y) + f(y) \quad \rho(x,a)(dy) \qquad \textit{gilt.}$$

5.1 Satz.

(*i*) *V ist exzessiv.*

(*ii*) *$(u_t + V(x_t))_{t \in T \cup \{0\}}$ ist ein Supermartingal.*

$$(iii) \qquad V(x) \geq \int r + V(x_2) \quad d\sigma_\pi[x]$$

$$\geq \int u_s + V(x_s) \quad d\sigma_\pi[x]$$

$$\geq \int u_{s'} + V(x_{s'}) \quad d\sigma_\pi[x]$$

$$\geq \lim_{t\in T} \int u_t + V(x_t) \quad d\sigma_\pi[x]$$

$$\geq \overline{\lim_{t\in T}} \int u_t \quad d\sigma_\pi[x]$$

für $s \leq s'$, $\quad s, s' \in T$.

BEWEIS: (ii) Für ein $\varepsilon > 0$ und jedes $x \in \mathcal{X}$ gibt es ein $\pi_x \in \Pi$ mit: $V(x) \leq \overline{\lim}_{t\in T} \int u_t \quad d\sigma_{\pi_x}[x] + \varepsilon$. Für Stoppzeiten s, s' $(s' > s)$ gilt wegen der Integral-Eigenschaften daher:

$$\int u_{s'[p_s]} + V(x_{s'[p_s]}) \quad d\sigma_\pi[p_s]$$

$$\leq \int u_{s'[p_s]} + \overline{\lim_{t\in T}} \int u_t \quad d\sigma_{\pi_{x_{s'[p_s]}}}[x_{s'[p_s]}] \quad d\sigma_\pi[p_s] + \varepsilon$$

$$= \overline{\lim_{t\in T}} \int u_t \quad d\sigma_{\pi^*}[p_s] + \varepsilon$$

$$\leq V(x_s) + \varepsilon,$$

wobei $\pi^* := (\pi^*)_{i\in\mathbb{N}}$ definiert ist durch:

$$\pi_i^*(x_0, a_0, ..., x_i)(\,) = \begin{cases} \pi_i(x_0, a_0, ..., x_i)(\,) & \text{für } i < s \\ \pi_{x_s,i}(x_s, a_s, ..., x_i)(\,) & \text{für } i \geq s \end{cases}$$

Daher gilt für s, s', $s \leq s'$, $s \in T \cup \{0\}$, und alle ε:

$$u_s + V(x_s) \geq u_s + \int u_{s'[p_s]} + V(x_{s'[p_s]}) \quad d\sigma_\pi[p_s] - \varepsilon.$$

Durch Grenzübergang in ε folgt die Supermartingaleigenschaft und für $s = 0$ folgt daraus (i).

(*iii*) folgt aus der Martingal-Eigenschaft von $(u_t + V(x_t))$, und die letzte Ungleichung von (*iii*) folgt aus Lemma 3.1(*v*) und $V(x_n) \geq \overline{\lim}_{t\in T} \int u_s \ \mathrm{d}\sigma_\pi[x_0, ..., x_n]$. ∎

Das obenstehende Lemma zeigt, daß ein Plan genau dann optimal ist, wenn überall Gleichheit vorliegt. Das führt zu den folgenden Definitionen:

DEFINITION.

(*i*) *Eine Aktion a ist* ε-**werterhaltend** *im Zustand x, wenn* $(u_1 = r)$ *gilt:*

$$V(x) \leq \int u_1(x_0, a_1, x_1) + V(x_2) \quad \rho(x, a)(dx_2) + \varepsilon$$

(*ii*) *Ein Plan* π *ist* ε-**werterhaltend** *im Zustand x, wenn gilt:*

$$V(x) \leq \lim_{t\in T} \int u_t + V(x_t) \quad \mathrm{d}\sigma_\pi[x] + \varepsilon$$

(*iii*) *Ein Plan* π *ist* ε-**egalisierend** *im Zustand x, wenn gilt:*

$$\overline{\lim_{t\in T}} \int u_t + V(x_t) \quad \mathrm{d}\sigma_\pi[x] - \varepsilon \leq \overline{\lim_{t\in T}} \int u_t \quad \mathrm{d}\sigma_\pi[x]$$

0-werterhaltende (0-egalisierende) Pläne werden **werterhaltend** (bzw. **egalisierend**) genannt. Aus den obigen Definitionen folgt, daß ein Plan 2ε-optimal ist, wenn er ε-werterhaltend und ε-egalisierend ist. Umgekehrt gilt, daß ein ε-optimaler Plan ε-werterhaltend und ε-egalisierend ist.

Der Begriff 'werterhaltend' könnte durch die Ausdrücke $V(x) = \lim_{n\in\mathbb{N}} \int u_n + V(x_n) \ \mathrm{d}\sigma_\pi[x]$, $V(x) = \lim \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[x]$ oder im Fall fast sicher endlicher Stoppzeiten durch $V(x) = \lim \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[x]$ definiert werden. Wie bekannt ist, reicht es zum Beweis der Optimalität von Plänen in diskontierten Modellen zu zeigen, daß ein Plan werterhaltend bezüglich des ersten Ausdrucks ist. Welche Definition von werterhaltend man wählt, kann von der Anwendung abhängig gemacht werden. Hier soll im Fall $T = T_e$ der zweite Ausdruck und im Fall $T = T_\infty$ der dritte Ausdruck zur Definition von werterhaltend genommen werden. Analog wird für den Begriff egalisierend verfahren. Um die Fragestellung weiter abzuklären, wird im folgenden untersucht, unter welchen Bedingungen die drei Ausdrücke gleich sind. Es wird sich zeigen, daß nicht generell $\lim_{n\in\mathbb{N}} \int u_n + V(x_n) \ \mathrm{d}\sigma_\pi[x] = \lim_{t\in T_e} \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[x] = \lim_{t\in T_\infty} \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[x]$ gilt.

Die Begriffe 'werterhaltende Aktion' und 'werterhaltender Plan' sind, wie ein Beispiel zeigen wird, nicht vollständig symmetrisch. So besteht zwar ein werterhaltender Plan fast sicher aus werterhaltenden Aktionen. Ein Plan aber, der nur werterhaltende Aktionen verwendet, braucht nicht werterhaltend bezüglich der Gleichung $V(x) = \lim_{t\in T_\infty} \int u_t + V(x_t)\ \mathrm{d}\sigma_\pi[x]$ zu sein. Mit Hilfe der transfiniten Induktion läßt sich lediglich $V(x) = \lim_{t\in T_e} \int u_t + V(x_t)\ \mathrm{d}\sigma_\pi[x]$ zeigen.

In den folgenden Lemmata wird die Darstellung der Funktion

$\overline{\lim}_{t\in T} \int u_t + V(x_t)\ \mathrm{d}\sigma_\pi[x]$ untersucht.

5.2 LEMMA. *Ist* $\overline{\lim}_{t\in T} \int u_t\ \mathrm{d}\sigma_\pi[x] > -\infty$, *dann gilt*

(*i*) $\lim_{n\in \boldsymbol{N}} u_n + V(x_n)$ *existiert und ist integrierbar.*

(*ii*) $$\lim_{t\in T_\infty} \int u_t + V(x_t)\ \mathrm{d}\sigma_\pi[x] = \int \lim_{n\in \boldsymbol{N}} u_n + V(x_n) > -\infty\ \mathrm{d}\sigma_\pi[x]$$

BEWEIS: (*i*) ergibt sich aus der folgenden Ungleichungskette:

$$\begin{aligned}
\int (u_n + V(x_n))^-\ \mathrm{d}\sigma_\pi[x] &\leq \int \left(u_n + \overline{\lim_{t\in T}} \int u_t\ \mathrm{d}\sigma_\pi[p_n]\right)^-\ \mathrm{d}\sigma_\pi[x] \\
&= \int \left(\overline{\lim_{t\in T}} \int u_{t+n} p_n\ \mathrm{d}\sigma_\pi[p_n]\right)^-\ \mathrm{d}\sigma_\pi[x] \\
&\leq \int \underline{\lim_{t\in T}} \int u^-_{t+n} p_n\ \mathrm{d}\sigma_\pi[p_n]\ \mathrm{d}\sigma_\pi[x] \\
&\leq \int\int u^-_{t'} p_n\ \mathrm{d}\sigma_\pi[p_n]\ \mathrm{d}\sigma_\pi[x] + \varepsilon \\
&\qquad \text{für ein } t' \\
&= \underline{\lim_{t\in T}} \int u^-_{t'}\ \mathrm{d}\sigma_\pi[x] + \varepsilon.
\end{aligned}$$

Die letzte Gleichung folgt aus Lemma 3.1. Der letzte Ausdruck ist nicht unendlich, da $\overline{\lim}_{t\in T} \int u_t\ \mathrm{d}\sigma_\pi[x] > -\infty$ und Bedingung A2 gilt, weshalb

$\int (u_n + V(x_n))^- \ d\sigma_\pi[x] < \infty$ für alle n gilt. Bedingung (i) folgt nun aus dem Martingalsatz (Bauer 1974, 60.1).

(ii) Wegen (i) gilt $\overline{\lim}_{n\in\boldsymbol{N}} u_n + V(x_n) = \underline{\lim}_{n\in\boldsymbol{N}} u_n + V(x_n)$ gilt. Satz 15.1 impliziert daher daher (ii). ∎

5.3 LEMMA. *Ist* $\overline{\lim}_{t\in T} \int u_t \ d\sigma_\pi[x] > -\infty$ *im Fall* $T = T_\infty$ *oder im Fall* $T = T_e$ *C2 erfüllt, so gilt*

(i) $\{(u_t + V(x_t))_{t\in T_\infty \cup \{0,\infty\}}\}$ *ist ein Supermartingal.*

(ii) $\{(u_t + V(x_t))^-_{t\in T_\infty}\}$ *ist gleichförmig integrierbar.*

BEWEIS: Fall $T = T_\infty$: Die Supermartingaleigenschaft von $(u_t + V(x_t))_{t\in T_\infty}$ wurde im Lemma 5.1 gezeigt. $\int \lim_{n\in\boldsymbol{N}} u_n + V(x_n) \ d\sigma_\pi[x] \neq -\infty$ gilt wegen des vorstehenden Lemmas. Die gleichförmige Integrierbarkeit folgt dann aus aus Proposition 3.1(4⇔2) von Engelbert-Engelbert (1979) und die Supermartingaleigenschaft ergibt sich aus Satz 15.5(ii).

Fall $T = T_e$: Wegen C2 (Satz 15.3(ii)) gilt:

$$\overline{\lim_{t\in T_e}} \int u_t \ d\sigma_\pi[x] \geq \overline{\lim_{t\in T_\infty}} \int u_t \ d\sigma_\pi[x] = \int \overline{\lim_{n\in\boldsymbol{N}}} u_n \ d\sigma_\pi[x]$$

für alle $x \in \mathcal{X}$, woraus

$$u_s + \overline{\lim_{t\in T}} \int u_t \ d\sigma_\pi[x] \geq u_s + \overline{\lim_{t\in T_\infty}} \int u_t \ d\sigma_\pi[x]$$

folgt.

Da $\left(u_s + \int \overline{\lim}_{t\in T} u_t \ d\sigma_\pi[x]\right)_{s\in T_\infty}$ gleichgradig integrierbar ist [ergibt sich aus Engelbert-Engelbert (1979, Proposition 3.1)], folgt: $\{(u_s + V(x_s))^-\}_{s\in T_\infty}$ ist gleichförmig integrierbar. Damit ist (ii) gezeigt. Mit Proposition 3.1 (2 ⇔ 4) von Engelbert-Engelbert (1979) folgt dann (i). ∎

5.4 LEMMA.

(i) *Ist* $\overline{\lim}_{t\in T} \int u_t \ d\sigma_\pi[x] > -\infty$, $\|V^+\| < \infty$ *und gilt C1, so folgt im Fall* $T = T_\infty$*:*

$$\lim_{t\in T_e} \int u_t + V(x_t) \ d\sigma_\pi[x] = \lim_{t\in T_\infty} \int u_t + V(x_t) \ d\sigma_\pi[x]$$

$$= \int \lim_{n\in\boldsymbol{N}} u_n + V(x_n) \ d\sigma_\pi[x]$$

(ii) *Im Fall* $T = T_e$ *gilt die Aussage, wenn zusätzlich C2 erfüllt ist.*

BEWEIS: Im Fall $T = T_\infty$ gilt wegen der Supermartingaleigenschaft:

$$\lim_{t\in T_e}\int u_t + V(x_t)\ \ \mathrm{d}\sigma_\pi[x] \geq \lim_{t\in T_\infty}\int u_t + V(x_t)\ \ \mathrm{d}\sigma_\pi[x]$$

Die Behauptung folgt nun aus Satz 15.5(i).

Im Fall $T = T_e$ fallen unter den Voraussetzungen C1 und C2 beide Modelle zusammen. ∎

Die obigen Lemmata gelten sowohl für Modelle mit endlichen Stoppzeiten, als auch für Modelle mit fast sicher endlichen Stoppzeiten. Sie implizieren aber nicht, daß $\lim_{t\in T_\infty}\int u_t + V(x_t)\ \ \mathrm{d}\sigma_\pi[x] \geq \overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[x]$ ist.

5.5 LEMMA. *Gilt* $\overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[x] > -\infty$ *und ist im Falle* $T = T_e$ *zusätzlich die Bedingung C1 erfüllt, dann gilt:*

$$\int \lim_{n\in\mathbb{N}} u_n + V(x_n)\ \ \mathrm{d}\sigma_\pi[x] \geq \overline{\lim_{t\in T}}\int u_t\ \ \mathrm{d}\sigma_\pi[x]$$

BEWEIS: Im Fall fast sicher endlicher Stoppzeiten folgt die Behauptung aus Lemma 5.1 und aus Lemma 5.4.

Der Fall endlicher Stoppzeiten: Wegen Satz 15.5(i) gilt:$\overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[p_m] \leq \int \overline{\lim}_{n\in\mathbb{N}} u_n\ \ \mathrm{d}\sigma_\pi[p_m]\ \forall_m$ und daher

$$\left\{u_m + \overline{\lim_{t\in T}}\int u_t\ \ \mathrm{d}\sigma_\pi[x_0\ldots x_m]\right\}^+ \leq \left\{u_m + \int \overline{\lim_{n\in\mathbb{N}}} u_n\ \ \mathrm{d}\sigma_\pi[x_0\ldots x_m]\right\}^+.$$

Die Folge der rechten Seiten der Ungleichung ist als Folge bedingter Erwartungswerte von $\overline{\lim}_{n\in\mathbb{N}} u_n$ gleichförmig integrierbar, woraus die gleichförmige Integrierbarkeit der Folge der linken Seiten folgt.

Wegen der gleichförmigen Integrierbarkeit ist folgende Abschätzung möglich und damit das Lemma gezeigt (Satz 15.5(i)):

$$\int \lim_{n\in\mathbb{N}} u_n + V(x_n)\ \ \mathrm{d}\sigma_\pi[x] \geq \int \lim_{n\in\mathbb{N}}\left\{u_n + \overline{\lim_{t\in T}}\int u_t\ \ \mathrm{d}\sigma_\pi[x_0\ldots x_n]\right\}\ \ \mathrm{d}\sigma_\pi[x]$$

$$\geq \lim_{n\in T}\int u_s + \overline{\lim_{t\in T}}\int u_t\ \ \mathrm{d}\sigma_\pi[x_0\ldots x_s]\ \ \mathrm{d}\sigma_\pi[x]$$

$$\geq \overline{\lim_{t\in T}}\int u_t\ \ \mathrm{d}\sigma_\pi[x].$$

Die letzte Ungleichung folgt aus Lemma 3.1(v). ∎

Am Ende dieses Paragraphen wird ein Beispiel angegeben, in dem C1 verletzt ist und die Ungleichung des vorstehenden Lemmas nicht gilt.

Gilt (A3"): $\overline{\lim}_{t\in T} \int u_t \ \mathrm{d}\sigma_\pi[x] > -\infty$ und sind im Falle $T = T_e$ zusätzlich die Bedingungen C1 und C2 erfüllt, und ist V nach oben beschränkt, dann gilt:

$$
\begin{aligned}
V(x) &\geq \lim_{n\in\mathbb{N}} \int u_n + V(x_n) \ \mathrm{d}\sigma_\pi[x] && (1)\\
&\geq \lim_{t\in T_e} \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[x] && (2)\\
&= \lim_{t\in T_\infty} \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[x] && (3)\\
&= \int \lim_{n\in\mathbb{N}} u_n + V(x_n) \ \mathrm{d}\sigma_\pi[x] && (4)\\
&\geq \overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_\pi[x] && (5)\\
&= \int \lim_{n\in\mathbb{N}} u_n \ \mathrm{d}\sigma_\pi[x] && (6)
\end{aligned}
$$

Im Fall endlicher Stoppzeiten gelten die Ungleichungen 1 und 2 und die Gleichheit in 4 schon unter Bedingung (A3). Die Ungleichung in 5 gilt unter Bedingung C1 und V^+ (s.a. Lemma 5.2). Um in 3 (bzw. 6) ' $\geq$' zu erreichen, müssen zusätzlich C1 und C2 gefordert werden. Ein Beispiel wird zeigen, daß die Bedingungen nicht ohne weiteres verletzt sein dürfen, wenn man die Existenz der Ungleichungskette erhalten will.

Im folgenden wird nun gezeigt, daß in den bisher untersuchten Modellen (negatives, diskontiertes, positives Modell)

$$\lim_{n\in\mathbb{N}} \int u_n + V(x_n) \ \mathrm{d}\sigma_\pi[x] = \lim_{t\in T} \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[x] \qquad \text{gilt.}$$

5.6 LEMMA. *Sei* $\overline{\lim}_{t\in T} \int u_t \ \mathrm{d}\sigma_\pi[x] > -\infty$ *und sind im Falle* $T = T_e$ *zusätzlich die Bedingungen C1 und C2 erfüllt.*

(*i*) *Ist* π *ein überall* ε*-optimaler Plan für ein* $\varepsilon > 0$*, dann gilt:*

$\{(u_n + V(x_n))^+_N\}$ *ist gleichförmig* $\sigma_\pi[x]$*-integrierbar und*

$$\lim_{n\in N}\int u_n + V(x_n)\ \ \mathrm{d}\sigma_\pi[x] \;=\; \lim_{t\in T_\infty}\int u_t + V(x_t)\ \ \mathrm{d}\sigma_\pi[x]$$

$$\geq\ \overline{\lim_{t\in T}}\int u_t\ \ \mathrm{d}\sigma_\pi[x]$$

(*ii*) *Die Voraussetzung in* (*i*) *ist erfüllt, wenn*

$$\left\|\left(V(.) - \inf_\pi \overline{\lim_{t\in T}}\int u_t\ \ \mathrm{d}\sigma_\pi[.]\right)^+\right\| < \infty \quad \textit{gilt.}$$

(*iii*) *Die Bedingung in* (*ii*) *ist richtig, wenn* $r \geq 0$ *und* V *nach oben beschränkt ist.*

BEWEIS: Im Fall fast sicher endlicher Stoppzeiten folgt die gleichförmige Integrierbarkeit von $\left(u_n + \overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[x_0...x_n]\right)^+_n$ aus der Konvergenz von $\left(u_n + \overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[x_0...x_n]\right)_n$ gegen $\overline{\lim}_{n\in N} u_n$; im Fall endlicher Stoppzeiten wie im Beweis des vorstehenden Lemmas 5.5. Wegen $u_n + \overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[x] + \varepsilon \;>\; u_n + V(x_n)$ ist $\{(u_n + V(x_n))^+\}_n$ gleichförmig integrierbar, und wegen Lemma 5.3 ist $\{(u_n + V(x_n))^-\}_n$ gleichförmig integrierbar. Aus der fast sicheren Konvergenz von $(u_n + V(x_n))_n$ und der gleichförmigen Integrierbarkeit folgt (cf. H. Bauer (1972, Satz 20.4)):

$$\lim_{n\in N}\int u_n + V(x_n)\ \ \mathrm{d}\sigma_\pi[x] \;=\; \int \lim_{n\in N} u_n + V(x_n)\ \ \mathrm{d}\sigma_\pi[x]$$

$$\geq\ \overline{\lim_{t\in T}}\int u_t\ \ \mathrm{d}\sigma_\pi[x].$$

Aus dem Lemma 5.5 folgt dann (*i*).

(*ii*) und (*iii*) folgen direkt. ∎

5.7 LEMMA. *Sei* $\overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[x] > -\infty$.

(*i*) *Ist* V *nach oben beschränkt und gelten C1 und C2, dann gilt:*

$\{(u_n + V(x_n))^+\}_n$ *ist gleichförmig integrierbar und*

$$\lim_{n\in N}\int u_n + V(x_n)\ \ \mathrm{d}\sigma_\pi[x] = \lim_{t\in T_\infty}\int u_t + V(x_t)\ \ \mathrm{d}\sigma_\pi[x]$$

$$\geq \overline{\lim_{t\in T}}\int u_t\ \ \mathrm{d}\sigma_\pi[x]$$

(*ii*) *Die Folgerung in* (*i*) *gilt, wenn* $\lim_{n\in\mathbb{N}} u_n \quad \sigma_\pi[x]$*-f.s. existiert,* V *nach oben beschränkt ist und* u_n *in* L^1*-Norm gegen* $\lim_{n\in\mathbb{N}} u_n$ *konvergiert.*

(*iii*) *Die Voraussetzung in* (*ii*) *ist richtig, wenn* $r \leq 0$.

BEWEIS: (*i*) $(u_n^+)_n$ ist gleichförmig integrierbar, und wegen $u_n + \|V^+\| > u_n + V(x_n)$ ist $\left\{(u_n + V(x_n))^+\right\}_n$ gleichförmig integrierbar (cf. H. Bauer (1972), 20.3).

(*ii*) und (*iii*) folgen direkt. ∎

Die beiden obigen Sätze zeigen, daß in den traditionellen Modellen die verschiedenen Definitionen von werterhaltend und egalisierend zusammenfallen. In dem hier vorgestellten allgemeinen Modell gilt das nicht.

Das folgende Beispiel zeigt, daß die Gleichung allein unter der Bedingung $\|V^+\| < \infty$ nicht notwendig richtig ist.

BEISPIEL. *Sei* X *die Menge der natürlichen Zahlen plus* $*$, $A = \{a, b\}$,
$\rho(n, b, \{*\}) = 1, \qquad \rho(n, a, x) = \frac{1}{2}\delta_{\{n+1\}}(x) + \frac{1}{2}\delta_{\{*\}}(x)$
$r(*, a, *) = 0, \quad r(n, a, x) = 2^{2n}\, 1_{\{n+1\}}(x) - 2^{2n}\, 1_{\{*\}}(x), \quad r(n, b, *) = 0$

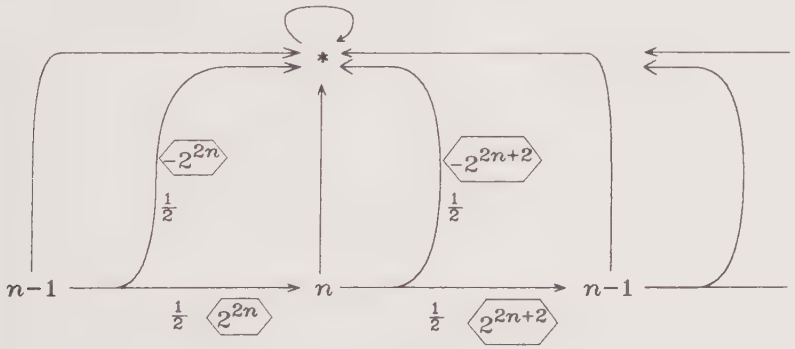

Bild 3

dann gilt $V(n) = 0$ *(s. Bild 3). Für den Plan* π*, der immer die Aktion* a *ergreift, gilt* $\int u_m + V(x_m) \; d\sigma_\pi[n] = \lim_{t\in T_e} \int u_t + V(x_t) \; d\sigma_\pi[n] = 0$ *und* $\overline{\lim}_{t\in T_\infty} \int u_t + V(x_t) \; d\sigma_\pi[n] = -\infty$. *Diese Werte ergeben sich unabhängig von der gewählten Menge der Stoppzeiten. Im Fall* $T = T_e$ *ist*

die Bedingung C1 verletzt, im Fall $T = T_\infty$ zusätzlich die Bedingung (A3), da $\overline{\lim}_{t\in T_e} \int u_t \ \mathrm{d}\sigma_\pi[n] = 0$ *und* $\overline{\lim}_{t\in T_\infty} \int u_t \ \mathrm{d}\sigma_\pi[n] = -\infty$.

Das vorstehende Beispiel zeigt, daß im Falle endlicher Stoppzeiten und im Fall fast sicher endlicher Stoppzeiten mit verschiedenen Nutzenfunktionen gearbeitet wird. Da man in Anwendungsproblemen davon ausgehen kann, daß der Prozeß irgendwann abbricht, zeigt daß Beispiel auch, daß es sinnvoll sein kann, mit endlichen Stoppzeiten zu arbeiten.

Im vorstehenden Beispiel ist die Bedingung A2 nicht erfüllt. Das folgende Beispiel zeigt, daß Gleichheit auch bei Gültigkeit der Bedingung A1+A2 nicht gelten muß.

BEISPIEL. *Ist X die Menge der natürlichen Zahlen vereinigt mit $*$, $A = \{a, b\}$,* $\rho(n, b, \{*\}) = 1$, $\rho(n, a, x) = \frac{1}{n}\delta_{\{n+1\}}(x) + (1 - \frac{1}{n})\delta_{\{*\}}(x)$
$r(*, ., *) = 0$, $r(n, ., x) = n! 1_{\{*\}}(x)$,

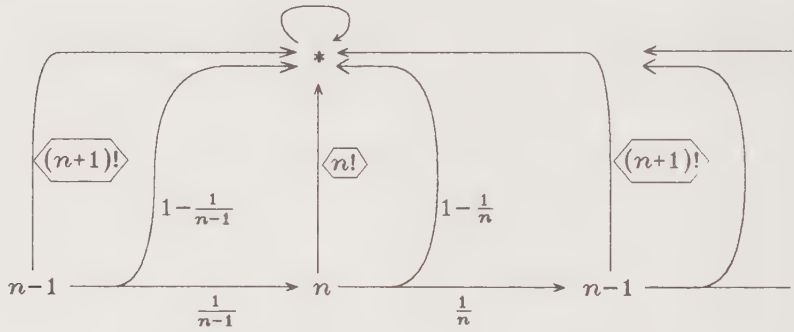

Bild 4

dann gilt $V(n) = n!$ (s. Bild 4). Für den Plan π, der immer die Aktion a ergreift, gilt: $\int u_m + V(x_m) \ \mathrm{d}\sigma_\pi[n] = \overline{\lim}_{t\in T_e} \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[n] = n!$ *und* $\overline{\lim}_{t\in T_\infty} \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[n] = 0$. *In diesem Fall ist Bedingung V^+ verletzt.*

Im Falle eines $M \subseteq M$ mit endlich additiven Maßen und bei gleichförmig beschränkter Folge $(u_t)_t$ gilt die Gleichung $\lim_{n\in N} \int u_n + V(x_n) \ \mathrm{d}\sigma_\pi[x] = \overline{\lim}_{t\in T_e} \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[x]$ nicht, wie Beispiele zeigen (cf. Dubins und Savage (1965)). Wegen der gleichförmigen Beschränktheit von $\{u_t + V(x_t)\}$

gilt aber $\lim_{t\in T_e} \int u_t + V(x_t)\ \mathrm{d}\sigma_\pi[x] = \overline{\lim}_{t\in T_\infty} \int u_t + V(x_t)\ \mathrm{d}\sigma_\pi[x] = \int \lim_{n\in N} u_n + V(x_n)\ \mathrm{d}\sigma_\pi[x]$.

6. Egalisierende Pläne.

In diesem Paragraph werden Bedinungen für egalisierende Pläne untersucht. Ferner wird gezeigt, daß in jedem $M \subset \mathcal{M}$ wenigstens ein egalisierender Plan existiert.

6.1 LEMMA. *Für jeden Plan π mit $\overline{\lim}_{t\in T} \int u_t\ \mathrm{d}\sigma_\pi[x] > -\infty$ gilt:*

(*i*) $$\overline{\lim_{t\in T}} \int V(x_t)\ \mathrm{d}\sigma_\pi[x] \geq \underline{\lim_{t\in T}} \int V(x_t)\ \mathrm{d}\sigma_\pi[x] \geq 0$$

(*ii*) *Ist C1 erfüllt oder V nach unten beschränkt, so gilt für π:*

$$\underline{\lim_{t\in T_e}} \int V(x_t)\ \mathrm{d}\sigma_\pi[x] \geq \underline{\lim_{t\in T_\infty}} \int V(x_t)\ \mathrm{d}\sigma_\pi[x]$$

BEWEIS: (*i*) Da $\int u_t + V(x_t)\ \mathrm{d}\sigma_\pi[x]$ in t konvergiert, gilt:

$$\overline{\lim_{t\in T}} \int u_t + V(x_t)\ \mathrm{d}\sigma_\pi[x] = \underline{\lim_{t\in T}} \int u_t + V(x_t)\ \mathrm{d}\sigma_\pi[x].$$

Da ferner

$$\underline{\lim_{t\in T}} \int u_t + V(x_t)\ \mathrm{d}\sigma_\pi[x] \geq \overline{\lim_{t\in T}} \int u_t\ \mathrm{d}\sigma_\pi[x],$$

folgt $$\underline{\lim_{t\in T}} \int V(x_t)\ \mathrm{d}\sigma_\pi[x] = \underline{\lim_{t\in T}} \int u_t + V(x_t)\ \mathrm{d}\sigma_\pi[x] - \overline{\lim_{t\in T}} \int u_t\ \mathrm{d}\sigma_\pi[x]$$

$$\geq 0.$$

Die Behauptung folgt nun aus den Regeln des $\underline{\lim}$.

(*ii*) Ist V nach unten beschränkt, so gilt $\underline{\lim}_{t\in T_e} \int V(x_t)\ \mathrm{d}\sigma_\pi[x] \geq \underline{\lim}_{t\in T_\infty} \int V(x_t)\ \mathrm{d}\sigma_\pi[x]$ (Satz 15.5(*ii*)).

Ist Bedingung C1 erfüllt, so gilt:

$$\lim_{t\in T_e}\int V(x_t)\ \ \mathrm{d}\sigma_\pi[x] = \lim_{t\in T_e}\int u_t + V(x_t)\ \ \mathrm{d}\sigma_\pi[x] - \int u_t\ \ \mathrm{d}\sigma_\pi[x]$$

$$= \lim_{t\in T_e}\int u_t + V(x_t) - \overline{\lim_{t\in T_e}}\int u_t\ \ \mathrm{d}\sigma_\pi[x]$$

wegen Rechenregeln des $\overline{\lim}_{t\in T}$

$$\geq \lim_{t\in T_\infty}\int u_t + V(x_t) - \overline{\lim_{t\in T_\infty}}\int u_t\ \ \mathrm{d}\sigma_\pi[x],$$

wegen C1 und Satz 15.5(i)

$$= \lim_{t\in T_\infty}\int V(x_t)\ \ \mathrm{d}\sigma_\pi[x],$$

womit (ii) bewiesen ist. ∎

Es sei angemerkt, daß das obige Lemma auch gilt, wenn $V(x_t)$ durch $\overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[x_0...x_t]$ ersetzt wird. Analog gilt das für andere Aussagen über den Ausdruck $V(x_t)$. Man kann nämlich zu einem Plan π ein $\mathcal{MM}$ definieren, in dem π der einzig mögliche Plan ist. Dazu geht man zu dem Zustandsraum $\mathcal{X}^*$ über und definiert die Menge der zulässigen Aktionen im Zustand $(x_0...x_t)$ durch $\Gamma'(x_0...x_t) := \{a_{x_0...x_t}\}$ mit $\pi_t(x_0...x_t)(\{a_{x_0...x_t}\}) = 1$. In dem neuen $\mathcal{MM}$ gilt dann $V'(x_0...x_t) = \overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[x_0...x_t]$.

Eine wichtige Eigenschaft egalisierender Pläne wird im folgende Lemma gezeigt.

6.2 LEMMA. *Für jeden Plan π mit $\overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[x] > -\infty$ gilt:*

(i) *π ist genau dann egalisierend, wenn* $\lim_{t\in T}\int V(x_t)\ \ \mathrm{d}\sigma_\pi[x] = 0$

(ii) *π ist genau dann ε-egalisierend, wenn* $\lim_{t\in T}\int V(x_t)\ \ \mathrm{d}\sigma_\pi[x] < \varepsilon$

BEWEIS: Für einen ε-egalisierenden Plan π gilt:

$$\lim_{t\in T}\int u_t + V(x_t)\ \ \mathrm{d}\sigma_\pi[x] - \overline{\lim_{t\in T}}\int u_t\ \ \mathrm{d}\sigma_\pi[x] < \varepsilon$$

Die Behauptung (ii) folgt nun auf Grund der Regeln des $\overline{\lim}$. (i) folgt analog. ∎

6.3 LEMMA. *Ist* $\int \overline{\lim}_{t\in T} u_t \ \mathrm{d}\sigma_\pi[x] > -\infty$, *so gilt*

(*i*) $\underline{\lim}_{n\in \boldsymbol{N}} V(x_n)$ *ist integrierbar bezüglich* $\sigma_\pi[x]$.

(*ii*) *Sind A3" und E1 erfüllt, so gilt*

$$\lim_{t\in T_\infty}\int V(x_t) \ \mathrm{d}\sigma_\pi[x] = \int \underline{\lim}_{n\in \boldsymbol{N}} V(x_n) \ \mathrm{d}\sigma_\pi[x].$$

(*iii*) *Ist im Fall* $T = T_e$ V *beschränkt, oder ist C1 erfüllt und* V *nach oben beschränkt, so gilt:*

$$\lim_{t\in T}\int V(x_t) \ \mathrm{d}\sigma_\pi[x] = \int \underline{\lim}_{n\in \boldsymbol{N}} V(x_n) \ \mathrm{d}\sigma_\pi[x].$$

BEWEIS: (*i*) $\lim_{n\in \boldsymbol{N}} u_n + V(x_n)$ ist integrierbar (Lemma 5.2). Wegen der Voraussetzung und der Sätze 15.3(*i*), 15.5(*i*) gilt

$$\infty > \overline{\lim_{t\in T}}\int u_t^+ \ \mathrm{d}\sigma_\pi[x] \geq \int \overline{\lim_{n\in \boldsymbol{N}}} u_n^+ \ \mathrm{d}\sigma[x]$$

und

$$\infty > \underline{\lim_{t\in T}}\int u_t^- \ \mathrm{d}\sigma_\pi[x] \geq \int \underline{\lim_{t\in T}} u_n^- \ \mathrm{d}\sigma[x].$$

$\overline{\lim}_{n\in \boldsymbol{N}} u_n$ ist daher integrierbar, woraus (*i*) folgt.

(*ii*) Ist $T = T_\infty$, so folgt die Behauptung aus Satz 15.2, da

$$\underline{\lim}_{n\in \boldsymbol{N}} V(x_n) = \lim_{n\in \boldsymbol{N}} (u_n + V(x_n)) - \overline{\lim}_{n\in \boldsymbol{N}} u_n \qquad \text{integrierbar ist}$$

und

$$\underline{\lim_{t\in T}}\int V(x_t) \ \mathrm{d}\sigma_\pi[x] = \lim_{t\in T}\int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[x] - \overline{\lim_{t\in T}}\int u_t \ \mathrm{d}\sigma_\pi[x]$$

$$> -\infty$$

gilt.

(*iii*) Ist $T = T_e$ und V beschränkt, so gilt:

$$\lim_{t\in T_e}\int V(x_t) \ \mathrm{d}\sigma_\pi[x] = \lim_{t\in T_\infty}\int V(x_t) \ \mathrm{d}\sigma_\pi[x]$$

$$= \int \underline{\lim}_{n\in \boldsymbol{N}} V(x_n) \ \mathrm{d}\sigma_\pi[x]$$

(Sätze 15.3(*ii*) und 15.5(*ii*)). Ist $T = T_e$, V nach oben beschränkt, so gilt:

$$\underline{\lim}_{t\in T_e}\int V(x_t)\ \ \mathrm{d}\sigma_\pi[x] \ \le\ \underline{\lim}_{t\in T_\infty}\int V(x_t)\ \ \mathrm{d}\sigma_\pi[x]$$

$$= \int \underline{\lim}_{n\in\mathbb{N}} V(x_n)\ \ \mathrm{d}\sigma_\pi[x]$$

(Satz 15.3(ii)). Aus dem vorstehenden Lemma folgt dann die Behauptung.∎

Das folgende Lemma zeigt nun, daß ein Plan Zustände, die negative Auszahlung liefern, in ferner Zukunft nicht mehr berührt.

6.4 LEMMA. *Ist* $\overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[x] > -\infty$ *und im Fall* $T = T_e$ V *nach oben beschränkt, so gilt*

(i) $\quad \sigma_\pi[x_0]\left(\overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[p_n] < -c\ ,\ \text{f.u.v.*}\ n\right) = 0 \quad \forall c > 0$

(ii) *Im Fall* $T = T_\infty$ *gilt:*

$$\sigma_\pi[x_0]\left(\overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[p_n] < c\ ,\ \text{f.u.v.}\ n\right) = 1 \quad \forall c > 0$$

(iii) *Ist im Fall* $T = T_e$ *zusätzlich zu den obigen Voraussetzungen:*

$$\overline{\lim}_{t\in T_e}\int \overline{\lim}_{s\in T}\int u_s\ \ \mathrm{d}\sigma_\pi[p_t]\ \ \mathrm{d}\sigma_\pi[p]$$

$$\ge \overline{\lim}_{t\in T_\infty}\int \overline{\lim}_{s\in T}\int u_s\ \ \mathrm{d}\sigma_\pi[p_s]\ \ \mathrm{d}\sigma_\pi[p]$$

für jede Vorgeschichte p, *die mit positiver Wahrscheinlichkeit durchlaufen wird, so gilt die Gleichheit in* (ii).

(iv) $\quad \sigma_\pi[x](V(x_n) < -c\ ,\ \text{f.u.v.}\ n) = 0 \quad \forall c > 0$

(v) $\quad \underline{\lim}_{t\in T}\sigma_\pi[x](V(x_t) < -c) = 0 \quad \forall c \in \mathbb{R}_+ - \{0\}$

(vi) $\quad \underline{\lim}_{t\in T_\infty}\int V^-(x_t)\ \ \mathrm{d}\sigma_\pi[x] = 0$

(vii) *Im Fall* $T = T_e$ *gilt unter der zusätzlichen Voraussetzung, daß* V *auch nach unten beschränkt ist:*

$$\underline{\lim}_{t\in T}\int V^-(x_t)\ \ \mathrm{d}\sigma_\pi[x] = 0$$

Die Bedingung $\overline{\lim}_{t\in T_-} \int u_t \ \mathrm{d}\sigma_\pi[x] \geq \overline{\lim}_{t\in T_\infty} \int u_t \ \mathrm{d}\sigma_\pi[x]$ in (iii) ist erfüllt, wenn $\overline{\lim}_{t\in T} \int u_t \ \mathrm{d}\sigma_\pi[p_n]$ gleichmäßig in (n, p_n) nach unten beschränkt ist.

BEWEIS: (i) und (ii) Fall $T = T_\infty$: Wegen des Martingal-Konvergenzsatzes gilt:

$$u_k + \int \overline{\lim_{n\in\mathbb{N}}} u_n \ \mathrm{d}\sigma_\pi[x_0 \cdots x_k] \longrightarrow \overline{\lim_{n\in\mathbb{N}}} u_n,$$

woraus die Behauptung auch für (ii) folgt.

Fall $T = T_e$: Wegen der Bemerkung im Anschluß an das erste Lemma gilt: $\underline{\lim}_{s\in T} \int \overline{\lim}_{t\in T} \int u_t \ \mathrm{d}\sigma_\pi[x_0 \ldots x_s] \ \mathrm{d}\sigma_\pi[x] \geq 0$. Ferner ist $\overline{\lim}_{t\in T} \int u_t \ \mathrm{d}\sigma_\pi[x] \leq V(x)$ nach oben beschränkt, wenn V nach oben beschränkt ist.

(i) Die mittlere Auszahlung bei Beobachtung des bedingten Plans $\pi[p_n]$ wird mit $I(\pi, p_n)$ abgekürzt:

$$I(\pi, p_n) := \overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_\pi[p_n].$$

Nimmt man nun an, daß die Behauptung nicht erfüllt ist, so gibt es eine Folge $(x_0, a_1, ..)$ [wie hier üblich $p_n = (x_0, a_1, ..., x_n)$] mit

$$\sigma_\pi[p_i]\,(\, I(\pi, p_{i+n}) < -c \ \text{ f.u.v. } n) \longrightarrow 1.$$

Wegen der Beschränktheit von V nach oben folgt:

$$\begin{aligned}
&\underline{\lim}_{s\in T} \int I(\pi, p_{s+i}) \ \mathrm{d}\sigma_\pi[p_i] \\
&\quad \leq \int \underline{\lim}_{n\in\mathbb{N}} I(\pi, p_{i+n}) \ \mathrm{d}\sigma_\pi[p_i] \qquad \text{(Satz 15.3}(ii)\text{)} \\
&\quad = \int_{[\underline{\lim}_{n\in\mathbb{N}} I(\pi, p_{i+n}) < -c]} \underline{\lim}_{n\in\mathbb{N}} I(\pi, p_{i+n}) \ \mathrm{d}\sigma_\pi[p_i] + \\
&\qquad \int_{[\underline{\lim}_{n\in\mathbb{N}} I(\pi, p_{n+i}) \geq -c]} \underline{\lim}_{n\in\mathbb{N}} I(\pi, p_{i+n}) \ \mathrm{d}\sigma_\pi[p_i] \\
&\quad \leq -c\, \sigma_\pi[p_i] \left(\underline{\lim}_{n\in\mathbb{N}} I(\pi, p_{i+n}) < -c \right) \\
&\qquad + \|V^+\|\, \sigma_\pi[p_i] \left(\underline{\lim}_{n\in\mathbb{N}} I(\pi, p_{i+n}) \geq -c \right)
\end{aligned}$$

Für genügend großes i ist der letzte Ausdruck daher echt kleiner 0. Wegen Teil (i) von Lemma 6.1 und der anschließenden Bemerkung gilt aber $\underline{\lim}_{s\in T}\int I(\pi, p_{s\oplus i})\ \ \mathrm{d}\sigma_\pi[x] \geq 0$ für alle x, π.

(iii) Nimmt man nun an, daß die Behauptung nicht erfüllt ist, so gibt es eine Folge $(x_0, a_1, ...)$ mit

$$\sigma_\pi[x_0, a_1..x_i]\,(\ I(\pi, p_n)\ \geq\ c\ \text{ für fast alle } n) \to 1.$$

Daher gilt:

$$\underline{\lim_{t\in T_e}}\int I(\pi,p_t)\ \ \mathrm{d}\sigma_\pi[x_0...x_i]\ \geq\ \underline{\lim_{t\in T_\infty}}\int I(\pi,p_t)\ \ \mathrm{d}\sigma_\pi[x_0...x_i]$$

$$= \int \underline{\lim_{n\in\mathbb{N}}}\, I(\pi,p_n)\ \ \mathrm{d}\sigma_\pi[x_0...x_i]$$

Die letzte Gleichung folgt aus Satz 15.2(ii). Die dortigen Voraussetzungen sind erfüllt, da mit V nach oben beschränkt $\underline{\lim}_{n\in\mathbb{N}}\ \ I(\pi,p_n)$ $\sigma_\pi[x]$-quasi-integrierbar ist und da wegen $\overline{\lim}_{s\in T}\int u_s + I(\pi,p_s)\ \ \mathrm{d}\sigma_\pi[x] - \overline{\lim}_{t\in T}\int u_t\ \ \mathrm{d}\sigma_\pi[x] = \overline{\lim}_{s\in T}\int I(\pi,p_s)\ \ \mathrm{d}\sigma_\pi[x]$ $\quad\overline{\lim}_{s\in T}\int I(\pi,p_s)\ \ \mathrm{d}\sigma_\pi[x]$ größer als $-\infty$ ist.

Weiter gilt:

$$\int \underline{\lim_{n\in\mathbb{N}}}\, I(\pi,p_n)\ \ \mathrm{d}\sigma_\pi[x_0...x_i]$$

$$= \int\limits_{[\underline{\lim}_{n\in\mathbb{N}}\, I(\pi,p_n)<-c]} \underline{\lim_{n\in\mathbb{N}}}\, I(\pi,p_n)\ \ \mathrm{d}\sigma_\pi[x_0...x_i]$$

$$+ \int\limits_{[\ \ c\leq\underline{\lim}_{n\in\mathbb{N}}\, I(\pi,p_n)<c]} \underline{\lim_{n\in\mathbb{N}}}\, I(\pi,p_n)\ \ \mathrm{d}\sigma_\pi[x_0...x_i]$$

$$+ \int\limits_{[\underline{\lim}_{n\in\mathbb{N}}\, I(\pi,p_n)\geq c]} \underline{\lim_{n\in\mathbb{N}}}\, I(\pi,p_n)\ \ \mathrm{d}\sigma_\pi[x_0...x_i]$$

$$\geq\ 0\ -\ c\,\sigma_\pi[x_0...x_i]\left(-c \leq \underline{\lim_{n\in\mathbb{N}}}\, I(\pi,p_n) < c\right)$$

$$+ c\sigma_\pi[x_0...x_i]\left(\underline{\lim_{n\in\mathbb{N}}}\, I(\pi,p_n) \geq -c\right)$$

Die letzte Abschätzung der Ungleichungskette folgt aus Teil (i). Für genügend großes i ist der letzte Ausdruck daher echt größer 0. Wegen Teil (i) von Lemma 6.1 und der anschließenden Bemerkung gilt aber $\underline{\lim}_{s\in T} \int I(\pi, p_s)\ \mathrm{d}\sigma_\pi[x] = \lim_{s\in T_f} \int u_s + I(\pi, p_s)\ \mathrm{d}\sigma_\pi[x] - \overline{\lim}_{t\in T} \int u_t\ \mathrm{d}\sigma_\pi[x] = 0$ für alle x, π.

(iv) folgt aus (i), da $\int \overline{\lim}_{n\in\mathbb{N}} u_n\ \mathrm{d}\sigma_\pi[x_0 \cdots x_k] \leq V(x_k)$ gilt.
(v) folgt aus (iv) und Paragraph 15a.
(vi) Aus (v) und Paragraph 15a folgt, daß

$$\overline{\lim_{t\in T_\infty}} \int V^-\ \mathrm{d}\sigma_\pi[x] = \int \overline{\lim_{t\in T}} V^-\ \mathrm{d}\sigma_\pi[x] = 0 \text{ gilt.}$$

Für den Fall der Beschränktheit von V nach unten ist die Aussage (vii) klar. ∎

Aus dem vorstehenden folgt nun, daß jeder egalisierende Plan die Menge $\{x \mid V(x) < \varepsilon\}$ unendlich oft berührt. Ferner ist jeder Plan, der diese Menge unendlich oft durchläuft, egalisierend.

6.5 SATZ.

(i) *Ist V beschränkt oder ist $T = T_\infty$ und E1 erfüllt, so gilt für einen ε-egalisierenden Plan $\pi \in \Pi$ ($\varepsilon > 0$, $\alpha > 0$)*

$$\sigma_\pi[x]\left(V(x_n) < \varepsilon\alpha \ \text{u.o.}*\right) \geq 1 - \alpha^{-1}.$$

(ii) *Unter den Voraussetzungen von (i) gilt für einen egalisierenden Plan $\pi \in \Pi$ und für alle $\delta > 0$,*

$$\sigma_\pi[x]\left(V(x_n) < \delta \ \text{u.o.}\right) = 1.$$

(iii) *Ist im Fall $T = T_\infty$ E1 und A3" erfüllt (bzw. im Fall $T = T_e$ $\|V^+\| < \infty$), dann ist jeder Plan π mit*

$$\sigma_\pi[x]\left(V(x_n) < \delta \ \text{u.o.}\right) = 1 \qquad \forall_\delta$$

egalisierend.

BEWEIS: (i) Für ε-egalisierende Pläne gilt: $\underline{\lim}_{t\in T} \int V(x_t)\ \mathrm{d}\sigma_\pi[x] \leq \varepsilon$. Aus Lemma 6.4$(v)$ folgt $\overline{\lim}_{t\in T} \int V^-(x_t)\ \mathrm{d}\sigma_\pi[x] = 0$, weshalb gilt:

---

* *unendlich oft*

$$\begin{aligned}
\varepsilon &\geq \varlimsup_{t\in T} \int V(x_t) \quad \mathrm{d}\sigma_\pi[x]. \\
&\geq \varlimsup_{t\in T} \int V^+(x_t) \quad \mathrm{d}\sigma_\pi[x] - \overline{\varlimsup_{t\in T}} \int V^-(x_t) \quad \mathrm{d}\sigma_\pi[x] \\
&\geq \varlimsup_{t\in T} \int\limits_{[V(x_t)\geq \varepsilon\alpha]} V^+(x_t) \quad \mathrm{d}\sigma_\pi[x] \\
&\geq \varlimsup_{t\in T} \sigma_\pi[x](V(x_t) \geq \varepsilon\alpha)\,\alpha\varepsilon, \qquad \text{woraus } (i) \text{ folgt.}
\end{aligned}$$

(ii) Ist π ein egalisierender Plan, dann folgt (i) für alle α und ε. Wähle nun zwei Folgen (α_ι), (ε_ι), mit $\alpha_\iota\varepsilon_\iota = \delta$ und $\alpha_\iota \to \infty$ und wende Teil (i) an.

(iii) Im Fall $T = T_\infty$ folgt die Behauptung aus der Gleichung

$$\varlimsup_{t\in T_\infty} \int V(x_t) \quad \mathrm{d}\sigma_\pi[x] = \int \varlimsup_{n\in\mathbb{N}} V(x_n) \quad \mathrm{d}\sigma_\pi[x].$$

Im Fall $T = T_e$ und nach oben beschränkter Auszahlungsfunktion V folgt die Behauptung aus der folgenden Ungleichungskette:

$$\begin{aligned}
0 &\leq \varlimsup_{t\in T_e} \int V(x_t) \quad \mathrm{d}\sigma_\pi[x] \leq \int \varlimsup_{n\in\mathbb{N}} V^+(x_n) \quad \mathrm{d}\sigma_\pi[x] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(Satz 15.3}(ii)\text{)} \\
&\leq \int\limits_{[\varlimsup_{n\in\mathbb{N}} V(x_n) \quad \mathrm{d}\sigma_\pi[x] > \delta]} \varlimsup_{n\in\mathbb{N}} V(x_n) \quad \mathrm{d}\sigma_\pi[x] + \delta = \delta.
\end{aligned}$$

Die Behauptung folgt nun durch Grenzübergang in δ. ∎

Aus Satz 6.5 folgt nun die Existenz eines egalisierenden Plans.

6.6 SATZ.

(i) *Ist A3" im Fall $T = T_\infty$ erfüllt (bzw. im Fall $T = T_e$ $\|V^+\| < \infty$), dann existiert ein egalisierender Plan.*

(ii) *Ist im Fall $T = T_\infty$ zusätzlich V nach oben beschränkt, dann kann der Plan aus (i) überall ε-optimal gewählt werden.*

BEWEIS: (i) Ist π_n ein $2^{-n}\ {}^1\varepsilon$-optimaler Plan ($n \geq 1$), dann folgt für $\alpha = 2$ und ein $t_n \in T$ aus dem vorstehenden Lemma (t_n kann ohne Beschränkung der Allgemeinheit aus T_e gewählt werden):

$$\sigma_{\pi_n}[x]\,(\,V(x_{t_n}) < 2^{-n}\varepsilon\,) \geq 4^{-1}$$

Für den zusammengesetzten Plan π' (vergl. Paragraph 1) folgt dann aus dem Satz von Borel-Cantelli:

$\sigma_{\pi'}[x]\,(V(x_i) < 2^{-n}\varepsilon$ f.u.v. $i\,) = 1$ Aus Satz 6.5 folgt dann, daß der Plan egalisierend ist.

(*ii*) Aus der $2^{-n-1}\varepsilon$-Optimalität der Pläne π_n folgt für den zusammengesetzten Plan π' und jede Vorgeschichte $p_{s_j^*}$ (Induktion über i):

$$V(x_{s_j^*}) \;\leq\; \int u_{s_i^*} + V(x_{s_i^*}) + \sum_{k=j}^{i} 2^{-k-1}\varepsilon \;\; \mathrm{d}\sigma_{\pi'}\left[p_{s_j^*}\right],$$

wobei $s_i^* = \bigoplus_k^i t_k$ ist.

Die Stoppzeiten s_n^* definieren eine Folge von σ-Algebren $\mathcal{A}_n := \sigma(s_1^*, ..., s_n^*)$ auf der Menge $(\mathcal{X} \times \mathcal{A})^\infty$. Für alle Pfade $(x_0...)$, jedes $n \in \mathbb{N}$ und jede Stoppzeit S auf der Folge von σ-Algebren $(\mathcal{A}_n)_n$ wird durch transfinite Induktion über die Struktur von S für alle j und alle Vorgeschichten $p_{s_j^*}$ gezeigt, daß gilt:

$$V(x_{s_j^*}) \leq \int u_{s_S^*} + V(x_{s_S^*}) + \sum_{k=j}^{S} 2^{-k-1}\varepsilon \;\; \mathrm{d}\sigma_{\pi'}\left[p_{s_j^*}\right].$$

Für die Stoppzeiten mit Struktur 0 (das sind genau die konstanten Stoppzeiten) ist die Aussage oben gezeigt worden.

Sei S eine Stoppzeit mit Struktur α und $p_{s_j^*}$ eine Vorgeschichte. Nehmen wir an, die Behauptung sei für alle Stoppzeiten mit kleinerer Struktur und j gezeigt, dann gilt:

$$\int u_{s_S^*} + V(x_{s_S^*}) + \sum_{k=j}^{S} 2^{-k-1}\varepsilon \;\; \mathrm{d}\sigma_{\pi'}\left[p_{s_j^*}\right]$$

$$\geq \int u_{s_{j+1}} + \int u_{s^*_{S\left[p_{s^*_{j+1}}\right]}} + V(x_{s^*_{S\left[p_{s^*_{j+1}}\right]}}) + \sum_{k=j+1}^{S\left[p_{s^*_{j+1}}\right]} 2^{-k-1}\varepsilon \;\; \mathrm{d}\sigma_{\pi'}\left[p_{s_{j+1}^*}\right]$$

$$+2^{-j-1}\varepsilon \;\; \mathrm{d}\sigma_{\pi'}\left[p_{s_j^*}\right]$$

$$\geq \int u_{s_{j+1}} + V(x_{s_{j+1}}) + 2^{-j-1}\varepsilon \;\; \mathrm{d}\sigma_{\pi'}\left[p_{s_j^*}\right]$$

$$\geq V(x_{s_j}).$$

Aus der Induktion und der Definition des $\overline{\lim}$ folgt nun für den Plan π':

$$V(x) \leq \overline{\lim_{t \in T_e}} \int u_t + V(x_t) \;\; \mathrm{d}\sigma_{\pi'}[x] + \varepsilon.$$

Das bedeutet im Fall $T = T_e$, daß der Plan ε-werterhaltend und damit ε-optimal ist.

Fall $T = T_\infty$: Ist V nach oben beschränkt, so gilt (Satz 15.5(i)):

$$\underline{\lim_{t \in T_e}} \int V(x_t) \;\; \mathrm{d}\sigma_{\pi'}[x] = \lim_{t \in T_\infty} \int V(x_t) \;\; \mathrm{d}\sigma_{\pi'}[x] = 0$$

weshalb zusammen mit der durch Induktion gezeigen Ungleichung gilt, daß der Plan ε-optimal ist. ∎

Man überlegt sich leicht, daß die Lemmata und Beweise auch in Modellen mit endlich additiven Maßen gelten. Die Forderung der gleichmäßigen Beschränktkeit von (u_t) versteht sich von selbst.

7. Überall ε-optimale Pläne.

In einigen Arbeiten wurde die Existenz überall ε-optimaler Pläne untersucht (cf. A. Hordijk (1976), H.-H. Bodewig (1979) und R.P. Kertz (1979)). A. Hordijk forderte in seiner Arbeit die absolute Konvergenz der Einschrittauszahlungen. R.P. Kertz (1979) wandte die Technik von A. Hordijk auf Modelle mit verschiebungsunabhängiger Funktion an. Insbesondere schließt sein Ergebnis das $\mathcal{GM}$ ein. H.-H. Bodewig übertrug die Methoden von A. Hordijk auf Modelle mit additiver Gewinnfunktion. Wir geben hier einen Beweis der Existenz überall ε-optimaler Pläne in diesem Modell.

Wir geben zuerst eine nützliche Abschätzung für die Menge der Pfade an, auf der ein Plan ε-optimal ist.

DEFINITION. *Ein Plan π ist* **überall ε-optimal in x_0 auf einer Menge vom Maß** δ, *wenn es eine Menge von Pfaden $\mathcal{H}$ mit $\sigma_\pi[x](\mathcal{H}) > \delta$ gibt, so daß für alle Pfade $(x_0, ..) \in \mathcal{H}$ und alle $n \in \mathbb{N}$ gilt:*

$$V(x_n) \leq \overline{\lim_{t \in T}} \int u_t \;\; \mathrm{d}\sigma_\pi[x_0, ..., x_n] + \varepsilon.$$

7.1 LEMMA. *Ist $\varepsilon \geq \delta > 0$ und π ein δ-optimaler Plan, dann ist π überall ε-optimal auf einer Menge vom Maß größer als $1 - \frac{\delta}{\varepsilon}$.*

BEWEIS: Aus der δ-Optimalität von π folgt: $\overline{\lim}_{t\in T} \int u_t \ \mathrm{d}\sigma_\pi[x] > -\infty$, weshalb

$$\int V(x_s) - \overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_\pi[x_0 \cdots x_s] \ \mathrm{d}\sigma_\pi[x] < \delta \qquad \forall_{s\in T} \qquad \text{gilt.}$$

Sei $\quad t_\varepsilon := \min\left\{ n \,\middle|\, V(x_n) > \overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_\pi[x_0 \cdots x_n] + \varepsilon \right\}, \qquad$ dann gilt

$$\int_{[t_\varepsilon < s]} V(x_{t_\varepsilon}) - \overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_\pi[x_0 \cdots x_{t_\varepsilon}] \ \mathrm{d}\sigma_\pi[x] < \delta \quad \forall_{s\in T},$$

woraus $\sigma_\pi[x](t_\varepsilon < s)\varepsilon < \delta$ folgt. Durch Grenzübergang in s folgt nun das Lemma. ∎

7.2 LEMMA. *Ist V beschränkt oder gilt $T = T_\infty$ und E1, so gibt es für jeden in $x \in \mathcal{X}$ wenigstens $\varsigma\delta$-optimalen Plan π mit ($\varsigma < 1, \quad \delta > 0$), eine Stoppzeit $t \in T$ mit:*

(*i*) $\quad \sigma_\pi[x](\quad V(x_i) \leq \delta, \quad \exists i \leq t \,) \geq 1 - 2\varsigma$

(*ii*) $\quad V(x_i) \leq \int u_{t[p_i]} + V(x_{t[p_i]}) \ \mathrm{d}\sigma_\pi[p_i] + \delta \quad \forall i \leq t.$

BEWEIS: Wegen A3 gilt für einen ε-optimalen Plan π $\overline{\lim}_{t\in T} \int u_t \ \mathrm{d}\sigma_\pi[x] > -\infty$. Sei $\varepsilon = \varsigma\delta$ und π ein ε-optimaler Plan und

$$s := \min\left\{ n \,\middle|\, V(x_n) > \overline{\lim_{t\in T}} \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[x_0 \cdots x_n] + \delta \right\}$$

$$\geq \min\left\{ n \,\middle|\, V(x_n) > \overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_\pi[x_0 \cdots x_n] + \delta \right\}$$

Da π ε-optimal ist, gilt wegen Lemma 7.1

$$\sigma_\pi[x](\quad s = \infty \quad) > 1 - \tfrac{\varsigma\delta}{\delta} = 1 - \varsigma \text{ und } \sigma_\pi[x](s \geq 1) = 1$$

Wegen Satz 6.5(*i*) (π ist ε-optimal und daher ε-egalisierend):

$$\sigma_\pi[x](\quad \exists_n \quad V(x_n) < \delta \quad) > 1 - \varsigma$$

$$\sigma_\pi[x](\quad \exists_n \quad V(x_n) < \delta, \quad s = \infty \quad) > 1 - 2\varsigma.$$

Für $t := t_{[V<\delta]} \wedge s$ gilt nun die Behauptung. $t_{[V<\delta]}$ ist die Ankunftzeit in $[V < \delta]$. ∎

Es wird nun die Existenz überall ε-optimaler Pläne gezeigt.

7.3 SATZ. *Ist V nach oben beschränkt, so existiert ein überall ε-optimaler Plan π.*

BEWEIS: Sei $(\varepsilon_i)_i$ eine Folge reeller positiver Zahlen mit $\sum_i \varepsilon_i \leq \frac{1}{2}\varepsilon$. In jedem Zustand x existiert ein ε_i-optimaler Plan π_i. Für diesen Plan π_i existiert eine Stoppzeit $t_i \in T$ mit:

$$\sigma_{\pi_i}[x]\left(\quad V(x_{t_i}) \leq \tfrac{1}{2}\varepsilon \quad\right) \geq \tfrac{1}{4}$$

$$V(x_0 \ldots x_j) \leq \int u_t + V(x_t) \;\; \mathrm{d}\sigma_{\pi_i}[x_0 \ldots x_j] + \frac{1}{2}\varepsilon_i \quad \forall j \leq t_i.$$

Sei π' der zusammengesetzte Plan (vergl.: Paragraph 1). Wir zeigen, daß dieser Plan werterhaltend und egalisierend ist. Den Zeitpunkt $\bigoplus_k^i t_k(p)$ bezeichnen wir als i-ten **Erneuerungspunkt** von π'. Wegen der Konstruktion existieren entlang eines Pfades unendlich viele Erneuerungen des Plans π' (im Fall $T = T_\infty$ gilt das nur fast sicher). Sei $e_n(h)$ die Zahl der Erneuerungen von π' entlang des Pfade h bis zum Zeitpunkt n.

Werterhaltend: Wie im Beweis von Satz 6.6 läßt sich mit transfiniter Induktion über die Stoppzeiten S, die zur Folge von σ-Algebren $\left(\sigma(t_1, t_1 \oplus t_2, \ldots, \bigoplus_1^i t_k)\right)_i$ adaptiert sind, und alle $p_i^* = p_{\bigoplus_k^i t_k}$ zeigen:

$$V(x) \leq \int u_{\bigoplus_i^S t_k} + V(x_{\bigoplus_i^S t_k}) + \sum_{k=i}^{e_S} \varepsilon_k \;\; \mathrm{d}\sigma_{\pi'}[p_i^*]$$

Aus der letzten Ungleichung folgt:

$$\overline{\lim_{t \in T_e}} \int u_t + V(x_t) \;\; \mathrm{d}\sigma_{\pi'}[x] \geq V(x) - \frac{1}{2}\varepsilon,$$

und daher für den Fall $T = T_e$, daß π $\frac{1}{2}$-werterhaltend ist. Für den Fall $T = T_\infty$ ist daher nur die Eigenschaft T_e-werterhaltend erfüllt.

Egalisierend: Wegen des Lemmas von Borel-Cantelli gilt: $\sigma_{\pi'}[x]\left(\quad V < \frac{1}{2}\varepsilon \quad u.o. \quad\right) = 1$. Nach Satz 6.5$(iii)$ folgt aus der Voraussetzung V nach oben beschränkt, daß π $\frac{1}{2}\varepsilon$-egalisierend ist (Fall $T = T_e$). Für den Fall $T = T_\infty$ folgt analog zum Beweis von Satz 6.5(iii) aus der Voraussetzung V nach oben beschränkt, daß $\overline{\lim}_{t \in T_e} \int V(x_t) \;\; \mathrm{d}\sigma_\pi[x] < \varepsilon$ ist. Aus beiden Teilen folgt nun die Behauptung. ∎

R.v. Dawen (1983) und Schäl-Sudderth (1984) geben einen Satz über die Existenz überall guter stationärer Pläne an, wenn in jedem Startpunkt und in jedem Submodell gute Pläne existieren. Eine einfache Transformation zeigt, daß dieses Ergebnis die Existenz überall optimaler Pläne impliziert. In dieser Arbeit wird die Annahme gemacht, daß $\int \sup_n u_n \ \mathrm{d}\sigma_{\pi'}[x] < \infty$ für jeden stationären Plan π gilt. Außerdem wird dort die Nutzenfunktion $\int \overline{\lim}_{t\in T} u_n \ \mathrm{d}\sigma_\pi[x]$ untersucht.

7.4 SATZ. *Sind X und A endlich, so existiert ein optimaler Plan.*

BEWEIS: Da X endlich ist, ist V nach oben beschränkt. Sei $\varepsilon > 0$ so gewählt, daß für alle nicht werterhaltenden Aktionen [d.h. $a \in \Gamma(x)$ mit $V(x) \neq \int u_1 + V(x_1) \ \rho(x, a|dx_1)]$ $V(x) > \int u_1 + V(x_1) \ \rho(x, a|dx_1) + \varepsilon$ gilt. Wegen der Endlichkeit von Aktionen- und Zustandsraum existiert ein solches ε. Sei $\varepsilon' < \varepsilon$, dann gilt für eine Aktion a mit $V(x) < \int u_1 + V(x_1) \quad \rho(x, a|dx_1) + \varepsilon'$, daß a werterhaltend ist. Wegen Satz 7.3 existiert ein überall $\frac{1}{n}\varepsilon$-optimaler Plan π_n für alle n. Da $V(x) = \sup_n (\overline{\lim}_{t\in T} \int u_t \ \mathrm{d}\sigma_{\pi_n}[x])$ für alle x gilt, kann man die zulässigen Aktionen auf die Aktionen einschränken, die in einem der Pläne π_n vorkommen, ohne daß sich die maximale mittlere Auszahlung ändert. Sie wird deshalb ebenfalls mit V bezeichnet. Jeder Plan π_n ist nur aus $\frac{1}{n}\varepsilon$-werterhaltenden Aktionen zusammengesetzt. Wegen der speziellen Wahl von ε sind diese Aktionen werterhaltend.

In dem neuen Modell gilt für jeden Plan π

$$V(x) = \overline{\lim_{t\in T_e}} \int u_t + V(x_t) \ \mathrm{d}\sigma_\pi[x]$$

(Beweis durch transfinite Induktion wie im Satz 6.6). Im Fall $T = T_e$ bedeutet das, daß der Plan π werterhaltend ist.

Aus Satz 6.6 folgt nun die Existenz eines egalisierenden Plans, der wegen der Endlichkeit von V die Gleichung

$$\underline{\lim_{t\in T_e}} \int V(x_t) \ \mathrm{d}\sigma_\pi[x] = \overline{\lim_{t\in T_\infty}} \int V(x_t) \ \mathrm{d}\sigma_\pi[x] = 0$$

erfüllt. Zusammen mit der Gleichung des letzten Abschnitts folgt, daß der Plan optimal ist. ∎

8. Werterhaltende Pläne.

In dieser Arbeit wird die Frage untersucht, wie die mittlere Auszahlung in dem hier vorgestellten $\mathcal{MCM}$ maximiert werden kann. Man könnte mit gleichem Recht die Frage nach der Minimierung der Auszahlungen, besser: Verluste, stellen. Als mittlere minimale bzw. duale Nutzenfunktion könnte man die Funktion:

$$\underline{V} := \inf_{\pi} \overline{\lim_{t\in T}} \int u_t \ \ \mathrm{d}\sigma_\pi[x]$$

wählen. Man sieht leicht, daß ein solches Modell dem $\underline{\lim}$-Modell verwandt ist, das durch folgende Auszahlungsfunktion gekennzeichnet ist (cf. W.D. Sudderth (1981)):

$$\overline{V} := \sup_{\pi} \underline{\lim_{t\in T}} \int -u_t \ \ \mathrm{d}\sigma_\pi[x].$$

Hierbei ist $\overline{V} = -\underline{V}$, und jede optimale Nutzenfunktion in dem einen Modell ist auch optimal in dem anderen.

Damit alle Größen eines solchen dualen Modells erklärt sind, bzw. die Integrale existieren, wird in diesem Abschnitt angenommen, daß

A3 $\quad \underline{V}(x) > -\infty \quad \forall_x$

gilt. Wegen $V \geq \underline{V}$; ist $u_s + V(x_s) \geq u_s + \underline{V}(x_s)$ und daher ist $u_s + \underline{V}(x_s)$ integrierbar für alle $s \in T$, $\pi \in \Pi$.

Eine wesentliche Martingaleigenschaft dieses Modells wird durch folgendes Lemma angegeben:

8.1 LEMMA. *Für jeden Plan* π *mit* $\overline{\lim_{t\in T}} \int u_t \ \ \mathrm{d}\sigma_\pi[x] > -\infty$ *gilt:*

(*i*) $\quad V \geq \underline{V}$,

(*ii*) $\quad \{u_t + \underline{V}(x_t)\}_{t\in T}$ *ist ein Submartingal,*

(*iii*) $\quad \lim_{n\in N} u_n + \underline{V}(x_n)$ *existiert.*

BEWEIS: (*ii*) Die Submartingaleigenschaft folgt, da gilt:

$$\int u_t + \underline{V}(x_t) \ \ \mathrm{d}\sigma_\pi[x_0 \ldots x_i] \ \geq \ \underline{V}(x_i).$$

(*iii*) Hierzu ist zu zeigen, daß $\sup_n \int (u_n + V(x_n))^+ \ \ \mathrm{d}\sigma_\pi[x] < \infty$ ist:

$$\int (u_n + \underline{V}(x_n))^+ \ \mathrm{d}\sigma_\pi[x] \leq \int (u_n + \overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_\pi[p_n])^+ \ \mathrm{d}\sigma_\pi[x]$$

$$\leq \overline{\lim_{t\in T}} \int \int u^+_{n+t} p_n \ \mathrm{d}\sigma_\pi[p_n] \ \mathrm{d}\sigma_\pi[x]$$

wegen Lemma 3.1 (vi)

$$\leq \overline{\lim_{t\in T}} \int u^+_{n+t} \ \mathrm{d}\sigma_\pi[x] + \varepsilon$$

Die letzte Abschätzung gilt für alle ε. Die Aussage folgt nun aus dem Martingal-Konvergenzsatz. ∎

Für die duale Nutzenfunktion sind die Untersuchungen aus den vorhergehenden Abschnitten übertragbar. Insbesondere lassen sich für dieses Modell Begriffe wie 'werterhaltend' und 'egalisierend' definieren. Die Funktion $\underline{V}$ hat andere Eigenschaften als V, so gilt $\underline{\lim}_{t\in T} \int \underline{V}(x_t) \ \mathrm{d}\sigma_\pi[x] \leq 0$, woraus $\sigma_\pi[x](\ \underline{V} \leq \delta \text{ u.o. }) = 1$ folgt [Anmerkung: $\underline{\lim}_{t\in T} \int V(x_i) \ \mathrm{d}\sigma_\pi[x] \geq 0$ impliziert $\sigma_\pi[x](\underline{V} \leq -\delta \text{ u.i. }) = 0$]. Man wird feststellen, daß der Begriff egalisierend etwas anders aufgefaßt werden muß: War im ursprünglichen Modell ein Plan egalisierend, so durchlief er die Menge $[V(x) < \delta]$ für beliebiges δ unendlich oft. Für die duale Nutzenfunktion gilt $\sigma_\pi[x](\ \underline{V} < -\delta \text{ u.o }) = 0$ (cf. W.D. Sudderth (1981)).

Auf eine Diskussion der Eigenschaften der dualen Nutzenfunktion wird verzichtet. Hier soll untersucht werden, wann eine Funktion werterhaltend ist. Es wird sich zeigen, daß das duale Modell zur Charakterisierung werterhaltender Pläne herangezogen werden kann. Betrachtet man nämlich die Menge $[V - \underline{V} = \vartheta \pm \varepsilon] \quad (\vartheta \in \mathbb{R})$, so ist ein Plan, der diese Menge nicht verläßt, werterhaltend.

8.2 LEMMA.

(i) $\{V(x_t) - \underline{V}(x_t)\}_{t\in T_\infty \cup \{\infty\}}$ *ist ein Submartingal.*

$$
\begin{aligned}
(ii) \qquad (V-\underline{V})(x) &\geq \lim_{n\in\mathbb{N}} \int (V-\underline{V})(x_n)\ \ \mathrm{d}\sigma_\pi[x] \\
&\geq \lim_{t\in T_\varepsilon} \int (V-\underline{V})(x_t)\ \ \mathrm{d}\sigma_\pi[x] \\
&\geq \lim_{t\in T_\infty} \int (V-\underline{V})(x_t)\ \ \mathrm{d}\sigma_\pi[x] \\
&= \int \lim_{n\in\mathbb{N}} (V-\underline{V})(x_n)\ \ \mathrm{d}\sigma_\pi[x] \\
&\geq 0
\end{aligned}
$$

BEWEIS: Die Supermartingaleigenschaft folgt aus der Supermartingaleigenschaft von $u_t + V(x_t)$ und der Submartingaleigenschaft von von $u_t + \underline{V}(x_t)$. Dabei folgt die Vertauschbarkeit von Limes und Integral aus Satz 15.5(ii). ∎

Aus der Supermartingaleigenschaft von $V - \underline{V}$ folgt, daß $V - \underline{V}(x_n)$ konvergiert. Der Prozeß $(V - \underline{V}(x_n))_n$ verändert sich daher in ferner Zukunft nicht sehr. Das nächste Lemma wird zeigen, daß das bedeutet, daß der Prozeß in ferner Zukunft werterhaltend ist.

8.3 LEMMA. *Ist $\varepsilon > 0$, $\delta > 0$, $\vartheta \in \mathbb{R}$ und existiert ein $x \in \mathcal{X}$ und ein Plan π mit:*

(i) $\qquad \vartheta + \varepsilon > (V - \underline{V})(x)$

(ii) $\qquad \sigma_\pi[x]\left((V-\underline{V})(x_n) > \vartheta - \varepsilon \quad \forall_{n>0}\right) > 1 - \delta,$

dann ist π; $2\varepsilon + \vartheta\delta$-werterhaltend in x. Ist $\delta = \frac{2\varepsilon}{\vartheta}$, dann π ist 4ε-werterhaltend.

BEWEIS: Aus der Ungleichungskette

$$
\begin{aligned}
\int V(x_t) - \underline{V}(x_t)\ \ \mathrm{d}\sigma_\pi[x] &\geq (\vartheta-\varepsilon)(1-\delta) \qquad ,(V-\underline{V}\geq 0) \\
&\geq (V-\underline{V}-2\varepsilon)(1-\delta) \\
&\geq V-\underline{V}-2\varepsilon-\delta\vartheta
\end{aligned}
$$

folgt für alle Stoppzeiten:

$$\int u_t + V(x_t)\ \ \mathrm{d}\sigma_\pi[x] \quad \geq \quad V - \underline{V} + \int u_t + \underline{V}(x_t)\ \ \mathrm{d}\sigma_\pi[x] - 2\varepsilon - \vartheta\delta$$

$$\geq \quad V - 2\varepsilon - \vartheta\delta$$

und daraus der erste Teil der Behauptung. ∎

8.4 SATZ. *Ist* $\varepsilon > 0$, $\delta > \delta' > 0$, $\vartheta \in \mathbb{R}$, *existiert ein* $x \in \mathcal{X}$ *und ein Plan* π *mit:*

(*i*) $\quad \vartheta + \varepsilon > (V - \underline{V})(x)$

(*ii*) $\quad \sigma_\pi[x]\left(\vartheta + \varepsilon > (V - \underline{V})(x_n) > \vartheta - \varepsilon \quad \forall_{n>0}\right) > 1 - \delta',$

dann ist π *überall* $2\varepsilon + \vartheta\delta$*-werterhaltend in* x *auf einer Menge vom Maß größer als* $1 - \frac{\delta'}{\delta}$.

BEWEIS: Sei

$$t_\varepsilon := \min\left\{ n > 0 \,\middle|\, \sigma_\pi[x_0 \ldots x_n]\,(\vartheta - \varepsilon < (V - \underline{V})(x_i) < \vartheta + \varepsilon \quad \forall_{i \geq 0}) \leq 1 - \delta \right\},$$

dann gilt:

$$\begin{aligned}
&\sigma_\pi[x]\,(t_\varepsilon < \infty) \\
&\quad = \sigma_\pi[x]\left(t_\varepsilon < \infty,\ \vartheta - \varepsilon < (V - \underline{V})(x_n) < \vartheta + \varepsilon \quad \forall_n\right) \\
&\qquad + \sigma_\pi[x]\left(t_\varepsilon < \infty,\ \vartheta - \varepsilon < (V - \underline{V})(x_n) < \vartheta + \varepsilon \quad \neg\forall_n\right) \\
&\quad \leq (1 - \delta)\sigma_\pi[x]\,(t_\varepsilon < \infty) + \delta',
\end{aligned}$$

wobei die Ungleichung aus der Voraussetzung folgt. Aus der Rechnung folgt:

$$\sigma_\pi[x]\,(t_\varepsilon = \infty) > 1 - \tfrac{\delta'}{\delta}.$$

Mit Wahrscheinlichkeit $1 - \frac{\delta'}{\delta}$ ist daher $\vartheta + \varepsilon > (V - \underline{V})(x_n)$ und $\sigma_\pi[x_0 \ldots x_n]\,(V - \underline{V} > \vartheta - \varepsilon\,\forall_{n \geq 0}) > 1 - \delta$ erfüllt. Aus Lemma 8.3 folgt dann die Behauptung. ∎

8.5 LEMMA. *Ist* $(\mathcal{O}_i)_i$ *eine offene Überdeckung der reellen Zahlen,* $(\delta_i)_i$ *eine Folge positiver reeller Zahlen,* π *ein Plan, ist weiter:*

$$\mathcal{H}_i^\pi := \left\{ h \,\middle|\, (V - \underline{V})(x_k) \in \mathcal{O}_i, \right.$$

$$\left. \sigma_\pi[p_k]\Big(\quad V - \underline{V}(x_n) \in \mathcal{O}_i \quad \forall_n \quad \Big) > 1 - \delta_i \; \exists_{k'} \quad \forall_{k>k'} \right\}$$

dann gilt:

(*i*) $\mathcal{H}_i^\pi \neq \emptyset$ *für ein* i.

(*ii*) $\sigma_\pi[x]\Big(\quad \bigcup_i \mathcal{H}_i^\pi \quad \Big) = 1.$

BEWEIS: Da gilt:

$$\left[\forall_{j>n} V - \underline{V} \in \mathcal{O}_i\right] \longrightarrow \left[\lim_{n\in\mathbb{N}} V - \underline{V} \in \mathcal{O}_i\right],$$

folgt $\sigma_\pi[x]$-fast-sicher (D. Revuz (1975, Proposition 2.2)):

$$\sigma_\pi[p_n]\,(\quad [\forall_{j\geq n} V - \underline{V}(x_j) \in \mathcal{O}_i \quad] \quad) \longrightarrow 1_{\{h \mid \lim V - \underline{V} \in \mathcal{O}_i\}}.$$

Daraus folgt, daß ein Pfad fast sicher in einer der Mengen $\mathcal{H}_i^\pi$ liegt. ∎

Betrachtet man die Menge:

$$\mathcal{G}_i := \left\{ h \,\middle|\, (V - \underline{V})(x_k) \in \mathcal{O}_i, \right.$$

$$\left. \sup_\pi \sigma_\pi[p_k]\Big(\quad (V - \underline{V})(x_n) \in \mathcal{O}_i \quad \forall_n \quad \Big) > 1 - \delta_i \quad \exists_k \right\},$$

so ist mit dieser Menge eine Stoppzeit $t_{\mathcal{G}_i}$ assoziiert, mit:

$$t_{\mathcal{G}_i}(x_0 \ldots) := \min\left\{ k \,\middle|\, (V - \underline{V})(x_k) \in \mathcal{O}_i, \right.$$

$$\left. \sup_\pi \sigma_\pi[p_k]\Big(\quad V - \underline{V}(x_n) \in \mathcal{O}_i \quad \forall_n \quad \Big) > 1 - \delta_i \right\},$$

Diese Stoppzeit ist selten endlich oder fast-sicher endlich, aber es gilt: $\bigwedge_i^\infty t_{\mathcal{G}_i} = t_{\bigcup_i^\infty \mathcal{G}_i} < \infty$ fast sicher für jeden Plan. Ab Erreichen dieser Stoppzeit ist

es möglich, den Prozeß mit großer Wahrscheinlichkeit in einer der Mengen $[V - \underline{V} \leq \vartheta_i \pm \varepsilon]$ zu halten. Daher ist der Prozeß ab diesem Zeitpunkt überall $2\varepsilon + \vartheta\delta$-werterhaltend ($\delta > \delta_i$) (s.Beweis des letzten Lemmas).

8.6 SATZ. *Sei* $\varepsilon > 0$, $(\vartheta_i)_{i \in \mathbf{N}}$ *eine steigende Folge reeller Zahlen,* $\vartheta_{i+1} - \vartheta_i \leq \frac{\varepsilon}{2}$, $\delta_i = \frac{\varepsilon}{4\vartheta_i}$ *und* π *ein Plan mit:* π *ist bis zum Zeitpunkt* $t \geq t_{\cup, \mathcal{G}}$, $\frac{\varepsilon}{2}$*-werterhaltend, dann ist* π *ein* ε*-werterhaltender Plan.*

BEWEIS: Das Satz folgt aus den den Sätzen 8.4 und 8.5. ∎

III Existenz optimaler Pläne

In den folgenden zwei Paragraphen soll die Existenz ε-optimaler stationärer Pläne gezeigt werden. Im verlaßbaren Fall existieren punktweise ε-optimale stationäre Pläne. Dieses Ergebnis ist eine Erweiterung einer Arbeit von W.D. Sudderth (1969). Im Paragraph 10 wird dann die Existenz von optimalen stationären Plänen bei endlichem Zustands- und Aktionenraum gezeigt.

9. Das verlaßbare Modell.

In diesem Abschnitt sollen die Ergebnisse auf verlaßbare Entscheidungsmodelle angewandt werden. Ein $\mathcal{MEM}$ ist **verlaßbar**, wenn

(i) für alle $x \in \mathcal{X}$ eine ausgezeichnete Aktion $*$ in $\Gamma(x)$ existiert mit $\rho(x,*) = \delta_x$ und

(ii) $r(x,*,x) = 0$.

Anschaulich bedeutet das, daß es in solchen Modellen erlaubt ist, den Prozeß zu verlassen, wann immer man will. Da man in verlaßbaren Modellen nur dann eineen Prozeß weiter verfolgen wird, wenn positiver Gewinn zu erwarten ist, ist dieses Modell mit den Modellen der positiven dynamischen Optimierung zu vergleichen. Aus diesem Grund gilt für verlaßbare Modelle $V \geq 0$.

Eine wichtige Eigenschaft der Funktion V ist, daß sie die kleinste positive exzessive Funktion im $\mathcal{MEM}$ ist.

9.1 Lemma. *Ist das $\mathcal{MEM}$ verlaßbar und ist $T = T_e$, so ist V die kleinste positive exzessive Funktion. Im Fall $T = T_\infty$ gilt das unter der zusätzlichen Bedingung C2.*

Beweis: Sei f eine positive exzessive Funktion, dann folgt für alle Stoppzeiten t durch transfinite Induktion:

$$f(x) \geq \int u_t + f(x_t) \;\; \mathrm{d}\sigma_\pi[x] \geq \int u_t \;\; \mathrm{d}\sigma_\pi[x].$$

Aus dieser Ungleichung folgt für alle Pläne π:

$$f(x) \geq \overline{\lim_{t \in T_e}} \int u_t \;\; \mathrm{d}\sigma_\pi[x].$$

Im Fall $T = T_e$ ergibt sich daraus direkt $f \geq V$. Im Fall $T = T_\infty$ folgt $f \geq V$ aus Satz 15.3(i). ∎

Verlaßbare Modelle haben die Eigenschaft, daß der unendliche Horizont durch den endlichen Horizont approximiert werden kann. Sei

$$U_n := \sup_\pi \int u_n \ \mathrm{d}\sigma_\pi[x]$$

die maximale mittlere Auszahlung bei Horizont n. Das folgende Lemma zeigt, daß U_n punktweise gegen V konvergiert.

9.2 LEMMA. *Ist das MçM verlaßbar und $T = T_e$ oder ist im Fall $T = T_\infty$ C2 erfüllt, dann gilt $U_i \uparrow V$.*

BEWEIS: Da $(r + U_n)_n$ eine monoton steigende Folge von Funktionen ist, gilt

$$\int r + U_n \ \mathrm{d}\sigma_\pi[x] \to \int r + \lim_{n \in N} U_n \ \mathrm{d}\sigma_\pi[x].$$

Für jedes ε existiert ein n mit

$$\int r + \lim_{n \in N} U_n \ \mathrm{d}\sigma_\pi[x] \leq \int r + U_n \ \mathrm{d}\sigma_\pi[x] + \varepsilon \leq U_{n+1} + \varepsilon.$$

Damit ist gezeigt, daß $\lim_{n \in N} U_n$ exzessiv ist. Da $\lim_{n \in N} U_n \leq V$ ist, folgt nun, daß $\lim_{n \in N} U_n = V$ ist. ∎

Es wird gezeigt, daß in verlaßbaren Modellen ε-optimale stationäre Pläne in jedem Startzustand x existieren.

9.3 SATZ (W.D. SUDDERTH (1969)). *Sei r eine nach oben beschränkte Funktion, n eine natürliche Zahl und $\varepsilon > 0$, so existiert für alle x ein stationärer Plan π mit*

$$\overline{\lim_{t \in T}} \int u_t \ \mathrm{d}\sigma_\pi[x] \geq U_n(x) - \varepsilon.$$

BEWEIS: (i) Sei ε_i eine Folge reeller Zahlen mit $0 = \varepsilon_0 < \varepsilon_1 < \ldots < \varepsilon_n = \varepsilon$ und sei

$$W(x) := \max_{0 \leq n} (U_k(x) - \varepsilon_k).$$

Wegen des endlichen Horizonts ist die Funktion W nach oben beschränkt und positiv.

Wähle einen stationären Plan π auf folgende Weise. Sei $\alpha = \frac{1}{2} \min_{1 \leq l \leq n-1} (\varepsilon_l - \varepsilon_{l-1})$ und $k(x)$ eine natürliche Zahl mit $W(x) = U_k(x) - \varepsilon_k$. Sei weiter a eine Auswahlfunktion mit:

$$a(x) = \begin{cases} * & , \text{ wenn } W(x) = 0 \\ g(x) & \text{sonst.} \end{cases}$$

$g(x)$ ist irgendeine in x erlaubte Aktion mit $\int r + U_{k-1} \ \mathrm{d}\rho(x, g(x)) \geq U_k(x) - \alpha$. Durch $\pi_n(x_0, \ldots, x_n) := a(x_n)$ wird ein Plan definiert. Es wird gezeigt, daß dieser Plan der gesuchte stationäre Plan ist.

Da W und r nach oben beschränkt sind, gilt:

$$\begin{aligned} \int r + W \ \mathrm{d}\rho(x, g(x)) &\geq \int r + U_{k-1} \ \mathrm{d}\rho(x, g(x)) - \varepsilon_{k-1} \\ &\geq U_k(x) - \alpha - \varepsilon_{k-1} \\ &\geq U_k(x) - \tfrac{1}{2}(\varepsilon_k - \varepsilon_{k-1}) - \varepsilon_{k-1} \\ &\geq W(x) + \tfrac{1}{2}(\varepsilon_k - \varepsilon_{k-1}) \\ &\geq W(x) + \alpha. \end{aligned}$$

9.4 LEMMA. *Sei* $\varepsilon_n(x_0 \cdots x_n) := r(x_{n-1}, a_{n-1}, x_n) + W(x_n) - W(x_{n-1}) \geq 0 \quad (n > 1)$, t *eine Stoppzeit, dann gilt für die Größen aus dem vorhergehenden Beweis:*

$$\begin{aligned} &\int u_t + W(x_t) \ \mathrm{d}\sigma_\pi[x] \\ &\quad = W(x) + \int \varepsilon_1(x_0, a_0, x_2) + \cdots + \varepsilon_{t(h)}(x_0 \cdots x_{t(h)}) \ \mathrm{d}\sigma_\pi[x] \end{aligned}$$

ε_i ist unabhängig von $(x_0 \cdots x_{i-2})$ und $\int \sum_i^t \varepsilon_i \ \mathrm{d}\sigma_\pi[x] \geq 0$.

BEWEIS: Da $\sum_{i-2}^{t(h)} \varepsilon_i = u_t + W(x_t) - W(x_0)$, $\int u_t^+ \ \mathrm{d}\sigma_\pi[x] < \infty$ und W nach oben beschränkt ist, ist $\sum^t \varepsilon_i$ integrierbar. Diese Aussage folgt durch transfinite Induktion über t. ∎

9.5 LEMMA. *Für* W *aus dem vorhergehenden Lemma gilt:*

$$\overline{\lim_{t \in T}} \int u_t \ \mathrm{d}\sigma_\pi[x] \geq \overline{\lim_{t \in T}} \int u_t + W(x_t) \ \mathrm{d}\sigma_\pi[x] \geq W(x).$$

BEWEIS: Ist $(\varepsilon_i)_{i \in \mathbf{N}}$ die Folge aus Lemma 9.4, dann folgt aus der Definition von π und ε_n

$$\int \varepsilon_i \ \mathrm{d}\sigma_\pi[x] \geq \alpha.$$

Weiter folgt aus dem Lemma 9.4:

$$\int u_t + W(x_t) \ \mathrm{d}\sigma_\pi[x] \geq W(x).$$

Für alle m gilt daher:

$$\begin{aligned} V(x) &\geq \int u_m + U_m(x_m) \ \mathrm{d}\sigma_\pi[x] - W(x) \\ &\geq \int u_m + W(x_m) \ \mathrm{d}\sigma_\pi[x] - W(x) \\ &= \int \varepsilon_1 + \cdots + \varepsilon_m \ \mathrm{d}\sigma_\pi[x] \\ &\geq \int_{[t_\delta \geq m]} m\alpha \ \mathrm{d}\sigma_\pi[x] \\ &= \sigma_\pi[x](t_\delta \geq m) m\alpha, \end{aligned}$$

wobei t_δ die Zeit des Verlassens des Prozesses ist.

Aus dem Vorstehenden folgt: $\lim_{n\in\boldsymbol{N}} \sigma_\pi[x](t_\delta \geq n) \leq \lim_{n\in\boldsymbol{N}} V(x) : n\alpha = 0$. Ist $\beta := \|W^+\|$, so gilt:

$$\begin{aligned} \int u_t + W(x_t) \ \mathrm{d}\sigma_\pi[x] - \int u_t \ \mathrm{d}\sigma_\pi[x] &= \int_{[t_\delta \geq t]} W(x_t) \ \mathrm{d}\sigma_\pi[x] \\ &\leq \beta\sigma_\pi[x](t_\delta \geq t) \end{aligned}$$

Die letzte Folge konvergiert gegen 0, womit das Lemma gezeigt ist. ∎

Mit dem Lemma 9.5 ist Satz 9.3 bewiesen. ∎

Die Existenz von ε-optimalen stationären Plänen folgt nun unter der Voraussetzung, daß der Zustandsraum endlich ist. Eine Beispiel bei J.v.d. Wal (1981) zeigt, daß dieses Ergebnis nicht auf abzählbaren Zustandsraum übertragbar ist.

9.6 SATZ. *Ist X endlich, r nach oben beschränkt und gilt im Fall $T = T_\infty$ C2, so existiert ein stationärer ε-optimaler Plan π.*

BEWEIS: Der Satz folgt, da bei endlichem Zustandsraum U_m gleichmäßig gegen V konvergiert. ∎

Im Satz 9.3 wird gefordert, daß r nach oben beschränkt ist. Ist man an stationären Plänen interessiert, so ist diese Forderung keine Einschränkung. Betrachtet man das Modell $(\mathcal{X}', \mathcal{A}', \Gamma', \rho', r')$ mit:

$$\mathcal{X}' = \mathcal{X} \cup \mathcal{X} \times \mathcal{A} \times \mathcal{X} \times \mathbb{N},$$

$$\mathcal{A} = \mathcal{A} \cup \{*\},$$

$$\Gamma'_{|\mathcal{X}} = \Gamma,$$

$$\Gamma'_{|\mathcal{X}\times\mathcal{A}\times\mathcal{X}\times\mathbb{N}} = \{*\},$$

$$\rho'(z) = \begin{cases} \sum_{n=0}^{\infty} \rho(x,a,y)\, \delta_{(x,a,y,n)}\, 1_{\{\lfloor |r(x,a,y)| \rfloor \geq n\}} & \\ & \text{für } z = (x,a) \\ \delta_{(x,a,y,n-1)} & z = ((x,a,y,n),*),\ n \geq 1 \\ \delta_y & z = ((x,a,y,0),*) \end{cases}$$

$$r'\big((x,a,n),*,(x,a,n-1)\big) = \begin{cases} \frac{r(x,a,y)}{\lceil r(x,a,y) \rceil} * & \\ 0 \qquad \text{sonst} & \end{cases}$$

so kann jeder stationäre Plan des einen Modells in einen des anderen Modells überführt werden, ohne daß sich die Auszahlung ändert. Das geänderte Modell hat den Vorteil einer durch 1 beschränkten Einschrittauszahlungsfunktion r.

Es sei angemerkt, daß dieser Satz auf positive Modelle (d.h. $r \geq 0$) übertragbar ist. Läßt man in einem positiven Modell zusätzlich zu, daß der Prozeß zu einem beliebigen Zeitpunkt gestoppt werden kann, so ist in beiden Modellen die mittlere maximale Auszahlungsfunktion V gleich. Ein

ε-optimaler Plan im gestoppten Modell kann aber beliebig im ursprünglichen Modell fortgesetzt werden, wobei die mittlere Auszahlung des fortgesetzten Plans größer ist als die des gestoppten Plans. Deshalb ist der im gestoppten Modell ε-optimale Plan auch im ursprünglich positiven Modell ε-optimal.

10. Stationäre optimale Pläne.

Bisher wurden Bedingungen für die Existenz ε-optimaler Pläne angegeben. Ferner wurde für den verlaßbaren Fall die Existenz von ε-optimalen stationären Plänen gezeigt. Mit Hilfe dieses Ergebnisses ist es nun möglich, die Existenz optimaler stationärer Pläne im Falle endlichen Zustands- und Aktionenraums zu beweisen. Der hier vorgestellte Beweis benutzt entscheidend die Technik der transfiniten Induktion auf der Menge der endlichen Stoppzeiten.

Zur Vorbereitung dieses Ergebnisses beweisen wir noch, daß stationäre optimale Pläne immer existieren, wenn der Zustandsraum endlich ist und in jedem Punkt ein optimaler Plan existiert.

10.1 LEMMA.

(*i*) *Ist X endlich und existiert ein optimaler Plan in x, dann existiert ein optimaler stationärer Plan in x.*

(*ii*) *Ist X endlich und existiert ein optimaler Plan, dann existiert ein optimaler stationärer Plan.*

BEWEIS: Ohne Beschränkung der Allgemeinheit kann angenommen werden, daß ein ausgezeichneter Zustand $*$ existiert, in dem nur eine Aktion a zulässig ist. Für diese Aktion a gelte $\rho(*, a)(\{x\}) > 0$ für alle Zustände.

Da ein optimaler Plan $\sigma_\pi[x]$-fast-sicher nur werterhaltende Aktionen benutzt, kann die Menge der zulässigen Aktionen in jedem $\Gamma(y)$ auf die Menge der werterhaltenden Aktionen eingeschränkt werden. In diesem geänderten Modell ist der Plan π zulässig, und daher ist die maximale mittlere Auszahlung für jeden Startpunkt x im neuen Modell unverändert.

In dem geänderten Modell gilt für jeden zulässigen Plan π (Beweis durch transfinite Induktion über $t \in T_e$):

$$V(x) = \overline{\lim_{t \in T_e}} \int u_t + V(x_t) \; \mathrm{d}\sigma_\pi[x].$$

Jeder Plan in dem geänderten Modell ist daher T_e- werterhaltend.

In dem geänderten Modell existiert ein egalisierender Plan π in jedem Zustand x des geänderten Modells. Jeder bedingte Plan $\pi[x_0 \ldots x_n]$ erreicht die Menge $\{x \,|V(x) \leq \delta\}$ mit Wahrscheinlichkeit 1 unendlich oft. Wegen der Endlichkeit des Zustandsraums kann δ so klein gewählt werden, daß gilt:

$$\{x \,|V(x) \leq \delta\} = \{x \,|V(x) \leq 0\}.$$

Satz 9.6 zeigt, daß ein stationärer Plan existiert, der diese Menge von jedem Zustand des eingeschränkten Modells mit Wahrscheinlichkeit größer $\frac{1}{4}$ erreicht. Aus dem Lemma von Borel-Cantelli folgt dann, daß diese Menge von dem stationären Plan mit Wahrscheinlichkeit 1 unendlich oft erreicht wird. Da der Zustandsraum endlich ist, ist V nach oben beschränkt und daher der Plan T_e-egalisierend.

Im Punkt $*$ existiert also ein stationärer optimaler Plan. Da von $*$ jeder Zustand mit positiver Wahrscheinlichkeit erreicht wird, ist der Plan auch optimal in allen anderen Zuständen des Modells. ∎

10.2 SATZ. *Sind $\mathcal{X}$ und $\mathcal{A}$ endlich, so existiert ein stationärer optimaler Plan.*

BEWEIS: Es existiert ein optimaler Plan π (Lemma 7.4). Wegen Lemma 10.1 existiert dann ein stationärer optimaler Plan. ∎

Beispiele bei Dubins und Savage (1965, Chap 3.9) zeigen, daß auf die Forderung der Endlichkeit von Zustands- und Aktionenraum nicht ohne weiteres verzichtet werden kann.

IV Optimale Markow Pläne

Bisher wurde die Existenz von guten stationären Plänen unter Endlichkeitsannahmen gezeigt. Wie Beispiele zeigen, existieren gute stationäre Pläne aber nicht, wenn der Aktionenraum abzählbar oder überabzählbar unendlich ist, cf. Beispiel 3.9.3 bei Dubins und Savage (1965). Deshalb soll in diesem Abschnitt untersucht werden, ob optimale Markow-Pläne existieren. Diese Untersuchungen sind eine Erweiterung einer Arbeit von Th.P. Hill (1979).

11. Egalisierende Markow-Pläne.

Oft wird man über das $\mathcal{MM}$ einige Informationen haben. Ist zum Beispiel die Funktion r negativ, oder ist r nach oben beschränkt und ein Diskontierungsfaktor $\beta < 1$ zu berücksichtigen, so ist jeder Plan egalisierend. Ist man in solchen Fällen an einem ε-optimalen Plan interessiert, so muß man nur Bedingungen für einen ε-werterhaltenden Plan untersuchen. Die Frage nach werterhaltenden Markow-Plänen soll hier nicht weiter untersucht werden, man überlegt sich leicht, daß ein Plan, der zum Zeitpunkt n eine ε_n-werterhaltende Aktion benutzt, selbst $\sum \varepsilon_n$-werterhaltend ist.

Existiert ein optimaler Plan, so kann die Menge der zulässigen Aktionen auf die Menge der werterhaltenden Aktionen eingeschränkt werden, ohne daß sich die maximale mittlere Auszahlung ändert. Deshalb wird hier die Frage untersucht, ob ein egalisierender Markow-Plan existiert. Ein solcher Plan wäre in dem geschilderten Fall optimal.

11.1 Satz. *Ist im Fall $T = T_e$ V nach oben beschränkt, oder sind im Fall $T = T_\infty$ E1 und A3" erfüllt, dann existiert ein Markow-Plan, der in allen Startzuständen x egalisierend ist.*

Beweis: Es wird ohne Beschränkung der Allgemeinheit angenommen, daß ein ausgezeichneter Zustand $*$ existiert, in dem nur eine ausgezeichnete Aktion a_* möglich ist. Für diesen Zustand $*$, jeden Zustand x und diese Aktion a_* gelte $\rho(*, a_*)(x) > 0$. Sei

$$\mathcal{G}_N := \{f \mid V(f) < N^{-1}\}.$$

Es wird gezeigt, daß in $*$ ein Markow-Plan existiert, der die Mengen $\mathcal{G}_N$ unendlich oft mit Wahrscheinlichkeit 1 durchläuft. Da dieser Plan jeden Zustand in der ersten Periode mit positiver Wahrscheinlichkeit erreicht, erreicht

jeder der bedingten Pläne $\pi[*, a_*, x]$ jede der Mengen $\mathcal{G}_N$ unendlich oft mit Wahrscheinlichkeit 1. Nach Satz 6.5(ii) ist der Plan $\pi' := \pi[*, a_*]$ in jedem Zustand $*$ egalisierend. Es braucht deshalb nur gezeigt werden, daß in $*$ ein egalisierender Markow-Plan existiert.

Wegen Satz 6.6 existiert ein egalisierender Plan π in $*$. Für diesen Plan gilt:

$$\sigma_\pi[*](t_{\mathcal{G}_N} < \infty) > 1 - \tfrac{1}{3N}.$$

$t_{\mathcal{G}_N}$ sei die Ankunftzeit in $\mathcal{G}_N$. Daher existiert eine natürliche Zahl n mit

$$\sigma_\pi[*](t_{\mathcal{G}_N} \leq n) > 1 - \tfrac{1}{2N}.$$

Sei $\left(\mathcal{X} \cup \{*\}, \mathcal{A} \cup \{a_*\}, \Gamma \cup \{(*, a_*)\}, \rho, r', n\right)$ das endliche Horizont-Modell mit:

(i) $r'(x, a, y) = 1_{\mathcal{G}_N}(y)$

(ii) Das Modell stoppt beim Erreichen der Menge $\mathcal{G}_N$, spätestens aber zum Zeitpunkt n.

Für die maximale mittlere Auszahlung $V_{N,n}$ gilt:

$$V_{N,n}(*) := \sup_\pi \int u'_n \; d\sigma_\pi[*] = \sup_\pi \; d\sigma_\pi[*](t_{\mathcal{G}_N} \leq n) \geq 1 - \tfrac{1}{2N}.$$

Es existiert ein Markow-Plan π mit:

$$\sup_\pi \sigma_\pi[*](t_{\mathcal{G}_N} \leq n) \geq 1 - \tfrac{1}{N}.$$

Das folgende Induktions-Argument vervollständigt den Beweis für festen Startpunkt $*$: Ist $(n_i)_{i \leq k}$ eine aufsteigende Folge natürlicher Zahlen und π ein werterhaltender Plan mit:

π ist Markow-Plan bis zum Zeitpunkt n_k,

π erreicht die Menge $\mathcal{G}_j$ zwischen den Zeitpunkten n_{j-1}, und n_j $(j \leq k)$ mit Wahrscheinlichkeit größer als $(1 - j^{-1})$,

dann läßt sich ein Markow-Plan für die nächste Periode konstruieren. In jedem Zustand x existiert ein Plan π_x, der die Menge $\mathcal{G}_{k+1}$ unendlich oft mit Wahrscheinlichkeit $(1 - \frac{1}{3(k+1)})$ erreicht. Benutzt man den Plan π im Startzustand * bis zum Zeitpunkt n_k und wählt dann einen Plan $\pi_{x_{n_k}}$, so erreicht man die Menge $\mathcal{G}_{k+1}$ unendlich oft mit Wahrscheinlichkeit $(1 - \frac{1}{3(k+1)})$. Es existiert eine natürliche Zahl n_{k+1}, so daß dieser Plan die Menge $\mathcal{G}_{k+1}$ zwischen den Zeitpunkten n_k und n_{k+1} mit Wahrscheinlichkeit $(1 - \frac{1}{2(k+1)})$

erreicht. Mit dem endlichen Horizont Argument aus dem zweiten Abschnitt kann nun ein Markow-Plan gefunden werden, der die Menge $\mathcal{G}_{k+1}$ zwischen den Zeitpunkten n_k und n_{k+1} mit Wahrscheinlichkeit $(1-(k+1)^{-1})$ erreicht.

Der so konstruierte Plan erreicht alle Mengen $\mathcal{G}_k$ mit Wahrscheinlichkeit 1. Aus Satz 6.5(*iii*) folgt, daß der Plan egalisierend ist. ∎

Aus Satz 11.1 ergibt sich, daß in einem Modell, in dem alle Aktionen werterhaltend sind, ein optimaler Markow-Plan existiert. Ferner zeigt der Beweis, daß es immer möglich ist, eine Menge mit einem Markow-Plan unendlich oft zu durchlaufen, wenn man sie mit irgendeinem Plan von jedem Zustand aus unendlich oft erreichen kann.

11.2 KOROLLAR. *Ist im $\mathcal{MCM}$ jede Aktion werterhaltend, ist V nach oben beschränkt und sind im Fall $T = T_\infty$ zusätzlich E1 und A3" erfüllt, dann existiert ein optimaler Markow-Plan.*

BEWEIS: Da jede Aktion werterhaltend ist, gilt für jeden Plan:

$$V(x) = \overline{\lim_{t \in T_e}} \int u_t + V(x_t) \;\; d\sigma_\pi[x]$$

Wegen Satz 11.1 gibt es einen egalisierenden Plan π. Für diesen Plan gilt wegen der Beschränktheit von V:

$$\underline{\lim_{t \in T_e}} \int V(x_t) \;\; d\sigma_\pi[x] = 0$$

und daraus folgt, daß der Plan optimal ist. ∎

12. Existenz ε-optimaler Markow-Pläne im Gambling Modell.

Bisher wurde festgestellt, daß egalisierende Pläne die Menge $[V = 0 \pm \delta]$ unendlich oft durchlaufen. Ferner ist es für werterhaltende Pläne hinreichend, eine Menge $[V - \underline{V} = c \pm \varepsilon]$ mit Wahrscheinlichkeit 1 unendlich oft zu durchlaufen. Bei der Frage nach guten Markow-Plänen im $\mathcal{MCM}$ kann man sich deshalb auf die Frage konzentrieren, ob es möglich ist, mit einem Markow-Plan in einer Menge zu bleiben und dabei vorbestimmte Punkte unendlich oft zu treffen. Diese Fragestellung wurde von Dubins und Savage mit ihrem $\mathcal{GM}$ aufgeworfen. Th.P. Hill (1979) untersuchte das $\mathcal{GM}$ für beliebigen Aktionenraum und endlichen Zustandsraum. Er konnte zeigen, daß in diesem Fall immer ein Markow-Plan existiert, der eine vorgegebene Menge

mit großer Wahrscheinlichkeit durchläuft. Hier sollen diese Ergebnisse weiter untersucht und die Ergebnisse von Th.P. Hill verallgemeinert werden.

Die Eigenschaft, eine Menge unendlich oft zu durchlaufen, führt zu folgender Definition.

DEFINITION.

(*i*) *Ein Zustand* $x \in \mathcal{X}$ *ist* **rekurrent**, *wenn für jedes* $\varepsilon > 0$ *ein Plan* π *mit* $\sigma_\pi[x](t_x < \infty) > 1 - \varepsilon$ *existiert.*

(*ii*) *Eine Menge* $\mathcal{C} \subset \mathcal{X}$ *ist* **rekurrent**, *wenn für alle* $x, y \in \mathcal{C}$ *und alle* $\varepsilon > 0$ *ein Plan* π *mit* $\sigma_\pi[x](t_y < \infty) > 1 - \varepsilon$ *existiert.*

(*iii*) *Eine Menge* $\mathcal{C} \subset \mathcal{X}$ *ist* **maximal rekurrent**, *wenn* $\mathcal{C}$ *rekurrent und maximal ist.*

Bei der Definition ist zu berücksichtigen, daß für alle Stoppzeiten $t \quad t > 0$ vorausgesetzt wird.

12.1 LEMMA. *Ist* ρ *ein Plan,* $\{\pi_k\}$ *eine Folge von Plänen,* $J \in \mathbb{N}$ *und* $\mathcal{C} \subset \mathcal{X}$ *eine maximale rekurrente Menge mit*

$$\sigma_\rho[x](\quad t_\mathcal{C} \oplus J < \infty \quad) =: \beta \leq 1,$$
$$\sigma_{\pi_k}[x](\quad t_\mathcal{C} < \infty \quad) \to 1 \quad \forall x \in \mathcal{C},$$

dann existiert für jedes $\varepsilon > 0$ *eine endliche Menge* $\mathcal{G} \subset \mathcal{C}^c$, $\mathcal{C}' := \{x_0, ..., x_n\} \subset \mathcal{C}$, *eine natürliche Zahl* $N \in \mathbb{N}$ *und eine Teilfolge* (π_l) *von* (π_k) *mit*

(*i*) $\sigma_\rho[x](\quad N > t_{\mathcal{C}'} \oplus J, \quad t_{\mathcal{G}^c} \oplus J > t_{\mathcal{C}'} \oplus J \quad) > \beta - \varepsilon,$

(*ii*) $\sigma_{\pi_l}[x](\quad t_\mathcal{G} > t_\mathcal{C} \quad) \to 1 \quad \forall x \in \mathcal{C}',$

BEWEIS: Für ein ε existiert eine minimale Zahl N mit

$$\sigma_\rho[x](\quad N > t_\mathcal{C} \oplus J, \quad) > \beta - \varepsilon,$$

dann existieren endliche Mengen $\mathcal{C}'$ und $\mathcal{G}$, die die Eigenschaft (*i*) erfüllen.

Angenommen, es existiert keine Folge σ_{π_l}, die (*ii*) erfüllt. Dann existiert ein $g \in \mathcal{G}$ mit

$$\sigma_{\pi_l}[x](\quad \infty > t'_\mathcal{C} > t_g \quad) > \varsigma > 0$$

für jedes ς. Daher ist g rekurrent, aber das ist ein Widerspruch zu der Annahme, daß $\mathcal{C}$ maximal rekurrent ist. ∎

12.2 LEMMA (s. Bild 5). *Sei* $\mathcal{D} \subset \mathcal{X}$, $\mathcal{C} \subset \mathcal{X}$ *eine maximal rekurrente Menge und* $\mathcal{D}_k := \left\{ x \in \mathcal{D} \,\middle|\, x \in \mathcal{C} \text{ oder } \inf_\pi \sigma_\pi[x](t_\mathcal{C} < \infty) > 1 - \frac{1}{k} \right\}$, *wobei* $\mathcal{D}_k \neq 0$. *Sei ferner* $\mathcal{B} \subset \mathcal{C}$ *und* π_k *eine Folge stationärer Pläne mit folgenden Eigenschaften:*

(α, i) *Es existiert ein* $y \in \mathcal{B}$, *so daß für ein* $\beta \in [0,1]$ *und jedes* k *gilt:*

$$\min\{m \,|\sigma_{\pi_k}[y](t_{D_k} \leq m) \geq \beta\}$$
$$= \min\{m \,|\sigma_{\pi_k}[x](t_{D_k} \leq m) \geq \beta, x \in \mathcal{B}\}$$

(α, ii) *Es existiert ein* $\zeta > 0$, *so daß für alle* $x_0, \ldots, x_n \in \mathcal{B}$, *das* y *aus* (α, i), *alle* $x_i \neq y$, *eine natürliche Zahl* N *und alle* k *gilt:*

$$\sigma_{\pi_k}[x_i](t_y < N) > \zeta,$$

(β) $\lim_k \sigma_{\pi_k}[x](t_{D_k} < \infty) = 1$ *gleichförmig in* $x \in \mathcal{B}$.

(γ) $\lim_k \sigma_{\pi_k}[x](t_B < \infty) = 1$ *gleichförmig in* $x \in \mathcal{C}$.

(δ) *Für jedes* $\varepsilon > 0$ *existiere ein Markow-Plan* μ *und natürliche Zahlen* N_1, N_2 *mit*

$$\sigma_\mu[x](t_\mathcal{B} \oplus N_1 < N_2, t_{(\mathcal{C}-\mathcal{B})} \oplus N_1 > N_2) > (1-\varepsilon)\sup_\pi \sigma_\pi[x](t_\mathcal{B} < \infty)$$

dann gilt: Es existiert eine Teilfolge π_l *von* π_k, *ein* $\zeta' > 0$ *und* $J_l \in \mathbb{N}$, *so daß für alle* $x_0, \ldots, x_n \in \mathcal{C}$, $N_* \in \mathbb{N}$ *gilt*

(i) $\sigma_{\pi_l}[x_i](t_{\mathcal{D}_l} < J_l - N_*) > \zeta'$,

(ii) $\lim_l \sigma_{\pi_l}[x_i](t_\mathcal{C} \oplus J_l < \infty) = 1 \quad \forall i$.

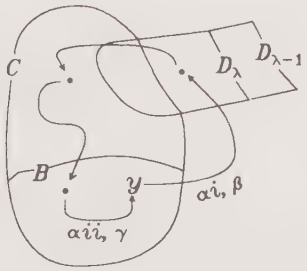

Bild 5

BEWEIS: (i) Sei π_k eine Folge von stationären Plänen mit $\sigma_{\pi_k}[x](t_{\mathcal{D}_k} < \infty) > 1 - \frac{1}{k}$ (Bedingungen β und γ). Wegen Bedingung γ kann die Folge π_k so reduziert werden, daß gilt:

$$\sigma_{\pi_k}[x](t_{\mathcal{B}} < \infty) > 1 - \tfrac{1}{k} \quad \forall_{x \in \mathcal{C} \setminus \mathcal{B}}.$$

Sei weiter $m_k \in \mathbb{N}$ eine Folge mit

$$\sigma_{\pi_k}[y](t_{\mathcal{D}_k} \leq m_k) \geq \beta > \sigma_{\pi_k}[x](t_{\mathcal{D}_k} < m_k) \quad \forall_{x \in \mathcal{B}}.$$

Eine solche Folge existiert wegen Bedingung (α, i).

(ii) Wähle in Bedingung (α, ii) eine natürliche Zahl N und eine Zahl k mit $k > 2$, so daß $\beta^k < \frac{\varepsilon}{4}$, $(1 - \frac{1}{k})^{3k(N+N_*)} > 1 - \frac{\varepsilon}{4}$, und $\frac{1}{k} < \frac{\varepsilon}{2}$ gilt. Sei ferner $m := m_k$, $J := N + m + N_*$ und $\pi = \pi_k$. Dann gilt die erste Ungleichung im Lemma für $\zeta' = \beta\zeta$.

Es bleibt zu zeigen, daß $\sigma_\pi[x](t_{\mathcal{C}} \oplus J \leq \infty) > 1 - \varepsilon$ für $x \in (x_0, ..x_n)$ ist. Sei s_1 die Ankunftzeit in $\mathcal{D}_k$, s_2 die Ankunftzeit in $\mathcal{C}$, s_3 die Ankunftzeit in $\mathcal{B}$ und $s = s_3 \oplus s_2 \oplus s_1$. Dann gilt:

$$\sigma_\pi[x](\, s^{N+N_*} < J\,) \leq \sigma_\pi[x](\, s \oplus N \oplus N_* - 1 < m + N + N_*)$$

$$= \sigma_\pi[x](s < m) < \beta \quad \forall x \in \mathcal{C}$$

und

$$\sigma_\pi[x](s^{k(N+N_*)} < J) \leq \beta^k < \tfrac{\varepsilon}{2}$$

Weiter gilt $\sigma_\pi[x](s < \infty) \geq (1 - \frac{1}{n})^3$ für alle $x \in \mathcal{C}$.

und daher $\sigma_\pi[x](s^{k(N+N_*)} < \infty) \geq (1 - \frac{1}{n})^{3k(N+N_*)}$ für alle $x \in \mathcal{C}$.

Aus dem letzten Ausdruck folgt:

$$\sigma_\pi[x](t_{\mathcal{C}} \oplus J < \infty) > 1 - \varepsilon. \quad \blacksquare$$

12.3 LEMMA. *Sind für eine rekurrente Menge $\mathcal{C} \subset \mathcal{X}$ und für $\mathcal{D} \subset \mathcal{C}$ die Bedingungen des Lemmas 12.2 erfüllt, dann existiert in jedem Startpunkt $x \in \mathcal{X}$ ein Markow-Plan μ mit*

$$\sigma_\mu[xp](\quad \mathcal{D} \text{ u.o. }) > \sup_\pi \sigma_\pi[xp](\quad t_{\mathcal{C}} < \infty \quad) - \varepsilon^* \qquad \forall_{p \in (A \times X)^*}.$$

BEWEIS: Der Beweis kann induktiv durchgeführt werden. Sei $(\varepsilon_i)_{i \in \mathbb{N}_0}$ eine Folge positiver reeller Zahlen mit $\sum_0^\infty \varepsilon_i \leq \varepsilon$. Angenommen es existiert

---

* $xp := (x, a_1, x_1, ..., x_n)$ für $p = (a_1, x_1, ..., x_n)$

ein Markow-Plan μ, eine Folge von Mengen $(\mathcal{G}_i)_{i\leq n}$ und natürliche Zahlen $(N_i)_{0\leq i\leq n}$ mit:

(i) $\quad \sigma_\mu[x]\,(\quad N_i > t_D \oplus N_{i-1} \leq t_C \oplus N_{i-1} \quad) > \zeta' \quad \forall_{i<n},$

(ii) $\quad \sigma_\mu[x]\,(\quad N_i > t_C \oplus N_{i-1} \quad)$

$$> \sup_\pi \sigma_\pi[x]\,(\quad t_C < \infty \quad) - \sum_j^i \varepsilon_j \quad \forall_{i<n}.$$

(iii) $\quad \sigma_\mu[x]\,(\quad \infty > t_C \oplus N_n \quad) > \sup_\pi \sigma_\pi[x]\,(\quad t_C < \infty \quad) - \sum_j^n \varepsilon_j.$

Für $N_0 = 0$ und einen stationären Plan μ mit:

$$\sigma_\mu[x]\,(\quad \infty > t_C \oplus N_0 \quad) > \sup_\pi \sigma_\pi[x]\,(\quad t_C < \infty \quad) - \varepsilon_0,$$

ist die Bedingung (iii) erfüllt und daher sind alle Bedingungen erfüllt.

Analog zu Lemma 12.1 kann man endliche Mengen $C' \subset C$, $\mathcal{G}_n \subset \mathcal{G}_{n+1} \subset C^c$ und einer natürliche Zahl N_{n+1} finden mit:

$$\sigma_\mu[x]\,(\quad t_{C'} < N'_{n+1},\ t_{\mathcal{G}^c} > t_C \quad)$$

$$> \sup_\pi \sigma_\pi[x]\,(\quad t_C < \infty \quad) - \sum_i^{n+1} \varepsilon_i.$$

Außerdem existiert eine Folge von Plänen π_l mit

$$\sigma_{\pi_l}[x]\,(\quad t_{\mathcal{G}} = \infty \quad t_C < \infty \quad) \to 1 \quad \forall x \in C'.$$

Sei, wie in Lemma 12.2, $N_* = N'_{n+1}$, μ irgendein Plan und $j_\mu = N_{n+1}$, so daß gilt:

(i) $\quad \sigma_{\pi_\mu}[x_i]\,(t_{D_\mu} < J_\mu - N_*) > \zeta',$

(ii) $\quad \sigma_{\pi_\mu}[x_i]\,(\quad t_{\mathcal{G}} = \infty, \quad t_C \oplus J_\mu < \infty) = 1 - \varepsilon$ für alle $x_i \in C$.

Wir definieren den Plan π folgendermaßen

bis zum Zeitpunkt $t_C \oplus N_n \cap N_{n+1}$ wählen wir den alten Plan μ,

danach wählen wir π_μ.

Der konstruierte Plan erfüllt für alle n die obigen Bedingungen. Eine Anwendung des Lemmas von Borel-Cantelli beendet den Beweis. ∎

Ein Beispiel wird zeigen, daß im Falle endlich additiver Maße die Bedingungen erfüllt sein können, aber daß trotzdem kein ε-optimaler Markow-Plan existiert.

12.4 LEMMA. *Ist $(C_i)_i$ eine Folge von maximal rekurrenten Mengen, $D' \subset \bigcup_i C_i$ und sind die Bedingungen von Lemma 12.2 für alle $D_i := D' \cap C_i$ und $C = C_i$ erfüllt, dann existiert in jedem Startzustand $x \in X$ ein Markow-Plan μ mit*

$$\sigma_\mu[xp](\quad D \text{ u.o. }) > \sup_\pi \sigma_\pi[xp](\quad t_{\bigcup_i C_i} < \infty \quad) - \varepsilon \qquad \forall_{p \in (A \times X)^*}$$

BEWEIS: Aus Lemma 12.3 folgt die Existenz eines guten Markow-Plans, der die Menge D mit Wahrscheinlichkeit $1 - \varepsilon$ für jedes $x \in C_i$ unendlich oft erreicht. Wir können annehmen, daß dieser Plan überall ε-optimal ist. Daher existiert für alle $x \in \bigcup_i C_i$ ein gemeinsamer guter Plan. Aus dem Lemma von W.D. Sudderth folgt nun die Existenz eines guten Plans in allen Zuständen x. ∎

12.5 LEMMA. *Ist der Zustandsraum X endlich, die Menge C maximal rekurrent und $D \subset C$, dann gibt es für alle ε und alle $x \in C$, einen Markow-Plan μ mit*

$$\sigma_\mu[x](D \text{ u.o. }) > 1 - \varepsilon \quad \forall x \in C.$$

BEWEIS: Angenommen, daß (π^n) eine Folge von stationären Plänen ist mit:

$$\sigma_{\pi^n}[x](\quad t_D < \infty, t_{C^c} \geq t_D \quad) > 1 - \varepsilon,$$

dann gilt $\pi_1^n[x](C) \to 1$. Da die Menge aller Maße auf C kompakt ist, existiert ein Häufungspunkt π_1 von π_1^n in der schwachen Topologie. Für den stationären Plan $\mu = (\pi_i)$ mit $\mu_i := \pi_1$ gilt:

$$\sigma_\mu[x](\quad C^\infty \quad) = 1.$$

Wegen der Endlichkeit von X existiert eine Menge von rekurrenten Zuständen $B \subset C$ analog zu π. Für die Folge π^n und die Mengen C und B sind die Bedingungen β, γ erfüllt.

Wegen der Rekurrenz der Menge B bezüglich μ ist die Bedingung δ erfüllt.

Die Bedingung (α, i) ist wegen der Endlichkeit von B erfüllt.

Die Bedingung (α, ii) ist erfüllt, da für jedes ς eine natürliche Zahl N existiert mit:

$$\sigma_\mu[x](\quad t_y < N \quad) > \varsigma. \blacksquare$$

12.6 LEMMA. *Ist der Zustandsraum $\mathcal{X}$ endlich, $\mathcal{D}$ eine Teilmenge des Zustandsraums, dann existiert für alle ε und alle $x \in \mathcal{C}$ ein Markow-Plan μ mit*

$$\sigma_\mu[x](\mathcal{D} \text{ u.o. }) > \sup_\pi \sigma_\pi[x](\quad \mathcal{D} \text{ u.o. }) - \varepsilon$$

BEWEIS: Wegen Lemma 12.5 existiert für alle maximal rekurrenten Mengen $\mathcal{C}$ mit $\mathcal{C} \cap \mathcal{D} \neq \emptyset$ ein guter Markow-Plan μ.

Im Falle $\sigma_\mu[x](\quad \mathcal{D} \text{ u.o. }) > 0$ existiert wenigstens eine maximal rekurrente Menge $\mathcal{C}$ mit $\mathcal{C} \cap \mathcal{D} \neq \emptyset$.

Für jeden anderen Zustand existiert ein stationärer Plan, der mit großer Wahrscheinlichkeit in eine der maximal rekurrenten Mengen führt. Der kombinierte Plan ist ein Markow-Plan und ε-optimal. ∎

Bisher wurde gezeigt, daß von jedem Punkt eines endlichem Zustandsraums eine vorgegebene Menge mit einem Markow-Plan unendlich oft mit hoher Wahrscheinlichkeit durchlaufen werden kann. Die folgenden Lemmata zeigen, daß dieser Markow-Plan unabhängig vom Startpunkt gewählt werden kann. Dazu wird zuerst gezeigt, daß es möglich ist, aus einer Menge ε-optimaler stationärer Pläne einen stationären Plan zu konstruieren, der überall ε-optimal ist. Dieser Satz ist eine einfache Version von Sätzen, die von W.D. Sudderth (1969), R.v. Dawen (1983) und Schäl-Sudderth (1984) gezeigt wurden.

12.7 LEMMA. *Ist $\mathcal{D}$ eine Teilmenge des Zustandsraums und existiert für alle ε und alle Zustände x und alle $M \in \mathfrak{M}$ $(\mathcal{X}, \mathcal{A}, \Gamma', \rho, r)$ mit $\Gamma'(x) \subset \Gamma(x)$ ein stationärer Plan π_x mit*

$$\sigma_{\pi_x}[x](\mathcal{D} \text{ u.o. }) > \sup_\pi \sigma_\pi[x](\quad \mathcal{D} \text{ u.o. }) - \varepsilon,$$

dann existiert ein stationärer Plan π mit:

$$\sigma_\pi[x](\mathcal{D} \text{ u.o. }) > \sup_\pi \sigma_\pi[x](\quad \mathcal{D} \text{ u.o. }) - \varepsilon, \ \forall_x \in \mathcal{X}.$$

BEWEIS: Sei $(\varepsilon_i)_i$ eine Folge positiv reeller Zahlen mit $\sum_1^\infty \varepsilon_i \leq \varepsilon$. Der Beweis erfolgt durch Konstruktion einer Folge von $M \in \mathfrak{M}$'en $(\mathcal{X}, \mathcal{A}, \Gamma^i, \rho, r)$ und einer aufsteigenden Folge von Mengen $\mathcal{G}^i \to \mathcal{X}$, so daß für alle i ein stationärer Plan π^i existiert mit

$$\sigma_{\pi^i}[x]\,(\ \mathcal{D}\text{ u.o. },(\mathcal{G}^i)^\infty\ )$$

$$> \sup_{\pi\in\mathcal{G}^{i+1}} \sigma_\pi[x]\,(\quad \mathcal{D}\text{ u.o. }) - \sum_{j=1}^{i}\varepsilon_j,\quad x\in\mathcal{G}^i.$$

π^{i+1} ist auf $\mathcal{G}^i$ gleich π^i.

Der Plan, der für $x\in\mathcal{G}^i$ die Aktion $\pi_1^i(x)$ ergreift, ist stationär und erfüllt die Behauptung des Lemmas.

Angenommen, ein $\mathcal{MEM}$ $(\mathcal{X},\mathcal{A},\Gamma^i,\rho,r)$, $\mathcal{G}^i$ und ein π mit der obigen Eigenschaft sind konstruiert. Sei $\Gamma^{i+1}(x)=\{\pi_1^i(x)\}$ für $x\in\mathcal{G}^i$. Dann existiert für ein $x\in(\mathcal{G}^i)^c$, $\varepsilon':=\frac{\varepsilon_{i+1}^2}{4}$ ein stationärer Plan π^{i+1} mit (s. Voraussetzung):

$$\sigma_{\pi^{i+1}}[x]\,(\mathcal{D}\text{ u.o. }) > \sup_\pi \sigma_\pi[x]\,(\quad \mathcal{D}\text{ u.o. }) - \varepsilon',\quad x\in\mathcal{G}^i.$$

Sei $\mathcal{G}^{i+1}:=\left\{x\;\middle|\;\sigma_{\pi^{i+1}}[x]\,(\mathcal{D}\text{ u.o. }) > \sup_{\pi\in\Gamma^{i+1}}\sup_\pi \sigma_\pi[x]\,(\quad \mathcal{D}\text{ u.o. }) - \frac{\varepsilon_{i+1}}{2}\right\}$,

dann gilt für alle $x\in\mathcal{G}^{i+1}$ nach Lemma 7.1

$$\sigma_{\pi^{i+1}}[x]\,(\ (\mathcal{G}^{i+1})^\infty\ ) > 1-\varepsilon'\frac{2}{\varepsilon_{i+1}}.$$

Sei t_i die Ankunftzeit in $\mathcal{G}^i$, dann gilt:

$$\sup_{\pi\in\Gamma}\sigma_\pi[x]\,(\quad \mathcal{D}\text{ u.o. })$$

$$= \sup_{\pi\in\Gamma}\int_{[\,t<\infty\,]} \sigma_\pi[p_t]\,(\quad \mathcal{D}\text{ u.o.}\quad)\ \mathrm{d}\sigma_\pi[x]$$

$$+\sigma_\pi[x]\,(\quad \mathcal{D}\text{ u.o.},\quad t=\infty\quad)$$

$$\le \sup_{\pi\in\Gamma}\int_{[\,t<\infty\,]} \sup_{\pi\in\Gamma}\sigma_\pi[p_t]\,(\quad \mathcal{D}\text{ u.o.}\quad)\ \mathrm{d}\sigma_\pi[x]$$

$$+\sigma_\pi[x]\,(\quad \mathcal{D}\text{ u.o.},\quad t=\infty\quad)$$

$$\le \sup_{\pi\in\Gamma}\int_{[\,t<\infty\,]} \sigma_{\pi^i}[p_t]\,(\quad \mathcal{D}\text{ u.o. },\quad (\mathcal{G}^i)^\infty\quad)\ \mathrm{d}\sigma_\pi[x]$$

$$+\ \sum^{i}\varepsilon_j\ +\ \sigma_\pi[x]\,(\quad \mathcal{D}\text{ u.o. },\quad t=\infty\quad)$$

$$= \sup_{\pi\in\Gamma}\int_{[\,t<\infty\,]} \sigma_{\pi^i}[p_t]\,(\quad \mathcal{D}\text{ u.o. },\quad (\mathcal{G}^i)^\infty\quad)\ \mathrm{d}\sigma_\pi[x]$$

$$+ \sum^{i} \varepsilon_j + \sigma_\pi[x](\quad \mathcal{D} \text{ u.o. }, \quad t = \infty \quad)$$

$$\leq \sup_{\pi \in \Gamma^{i+1}} \sigma_\pi[x](\quad \mathcal{D} \text{ u.o.} \quad) + \sum^{i} \varepsilon_j$$

$$\leq \sigma_{\pi^{i+1}}[x](\quad \mathcal{D} \text{ u.o. }, \quad \mathcal{G}^{i+1} \quad) \, d\sigma_\pi[x] + \sum^{i+1} \varepsilon_j,$$

für alle $x \in \mathcal{G}^{i+1}$, womit das Lemma bewiesen ist. ∎

12.8 SATZ. *Ist $\mathcal{X}$ endlich, $\mathcal{D}$ eine Teilmenge des Zustandsraums und $\varepsilon > 0$, dann existiert ein Markow-Plan μ mit*

$$\sigma_\mu[x](\mathcal{D} \text{ u.o. }) > \sup_\pi \sigma_\pi[x](\quad \mathcal{D} \text{ u.o. }) - \varepsilon, \forall x \in \mathcal{X}.$$

BEWEIS: Ohne Beschränkung der Allgemeinheit wird angenommen, daß jeder Zustand x die Form $(x_1, n) \quad$ mit $n \in \mathbb{N}$ hat und daß von einem Zustand (x_1, n) ein Zustand (y_1, m) nach genau $n - m$ Schritten erreicht werden kann. Unter dieser Voraussetzung ist jeder Markow-Plan stationär. Wegen Satz 12.5 existiert in jedem Punkt x ein stationärer Plan π_x mit

$$\sigma_{\pi_x}[x](\mathcal{D} \text{ u.o. }) > \sup_\pi \sigma_\pi[x](\quad \mathcal{D} \text{ u.o. }) - \varepsilon.$$

Wegen Lemma 12.7 gibt es einen gleichmäßig stationären Plan. ∎

13. Existenz eines ε-optimalen Markow-Plans im $\mathcal{MCM}$.

In diesem Abschnitt sollen die Ergebnisse auf das $\mathcal{MCM}$ übertragen werden. Wie der Beweis zeigt, existieren ε-optimale Markow-Pläne im $\mathcal{MCM}$, wenn für jedes $\mathcal{GM}$ $(\mathcal{X}, \mathcal{A}, \Gamma, \rho, 1_y)$ mit $y \subset \mathcal{X}$ ein ε-optimaler Markow-Plan existiert. Das ist insbesondere der Fall, wenn $\mathcal{X}$ endlich ist. Wir beschränken uns beim Existenzbeweis ε-optimaler Markow-Pläne im $\mathcal{MCM}$auf den Fall endlichen Zustandsraums.

13.1 SATZ. *Ist $\mathcal{X}$ endlich und sind im Fall $T = T_\infty$ C2 und A3" erfüllt, dann existiert ein überall ε-optimaler Markow-Plan für das $\mathcal{MCM}$.*

BEWEIS: Da $\mathcal{X}$ endlich ist, kann $\|V^+\| = 1$ angenommen werden. Sei $\varepsilon' = \frac{\varepsilon}{16}$, $(\vartheta_i)_i$ mit $\vartheta_1 = 1$ $\vartheta_{i+1} = \vartheta_i - \varepsilon'$ und $\delta = \frac{\varepsilon}{4}$. Sei ferner π ein überall ε'-optimaler Plan. π kann so gewählt werden, daß $\sigma_\pi[x_0 \ldots x_j](V < \varepsilon' \, u.o.) = 1 \quad \forall_j$ gilt (cf. Satz 6.6(ii), Satz 6.5(ii)).

Für wenigstens ein i gilt: $\mathcal{H}_i^\pi \neq \emptyset$ (s. Paragraph 8). Sei nun

$$\mathcal{G}_i := \left\{ x \,\middle|\, \sup_\pi \sigma_\pi [x] \left(\vartheta + \varepsilon' > V - \underline{V} > \vartheta - \varepsilon' \right) > 1 - \delta_i \right\}.$$

Da der Zustandsraum endlich ist, können die Werte δ_i so gewählt werden, daß gilt

$$\sup_\pi \sigma_\pi [x] \left(\vartheta + \varepsilon' > V - \underline{V} > \vartheta - \varepsilon' \right) = 1 \quad \forall_{\mathcal{G}_i}.$$

Sei $(\mathcal{X}, \mathcal{A}, \Gamma^i, \rho^i, r^i)$ ein $\mathcal{GM}$ mit folgender Eigenschaft: Alle Punkte außerhalb von $\mathcal{G}_i$ sind absorbierend und innerhalb von $\mathcal{G}_i$ sind alle Aktionen wie im ursprünglichen Modell erlaubt. Die Nutzenfunktion ist $r^i = 1_{\{ V(x) < \varepsilon' \}}$. Da der Zustandsraum endlich ist, existiert in diesem Modell ein Markow-Plan π^i mit (Satz 12.8):

$$\sigma_{\pi^i} [x_0, ..., x_j] \Big(\quad x_n \in \mathcal{G}_i \, \forall_n, \quad V(x_m) < \varepsilon' \text{ f.u.v. } m \Big)$$
$$> 1 - \delta \qquad \forall_j \quad \forall_{x_j \in \mathcal{G}_i}.$$

Für jeden Plan π im ursprünglichen Modell, der zum Zeitpunkt n die Aktion $\pi_n^i(x_0, ..., x_n)$ auswählt ($x_n \in \mathcal{G}_i \, \forall_n$), gilt:

$$\sigma_\pi [x_0, ..., x_j] \Big(\quad \vartheta_i - \varepsilon' < (V - \underline{V})(x_n) < \vartheta_i + \varepsilon' \quad \forall_n, \quad V(x_m) < \varepsilon' \text{ u.o.} \quad \Big)$$
$$> 1 - \delta \quad \forall_j \quad x_j \in \mathcal{G}_i.$$

Alle bedingten Pläne $\pi^i [p, x]$ sind $2\varepsilon' + \vartheta_i \delta \leq \frac{\varepsilon}{4}$-werterhaltend für alle $x \in \mathcal{G}_i$, $p \in \mathcal{X}^*$ (Lemma 8.2(i)) und wegen der Beschränktheit von V durch 1 gilt:

$$\underline{\lim_{t \in T}} \int V(x_t) \; \mathrm{d}\sigma_\pi [p, x] \leq \varepsilon'(1 - \delta) + \delta < \frac{\varepsilon}{4}.$$

π ist daher $\frac{\varepsilon}{4}$-egalisierend für alle Pfade $p \in \mathcal{X}^*$ und $x \in \mathcal{G}_i$. Für jedes i ist daher jeder bedingte Plan $\pi [x_0, ..., x_n]$ mit $\pi_j(x_0, ..., x_j) = \pi_j^i(x_0, ..., x_j)$ $\forall x_j \in \mathcal{G}_i, j \in I\!N$, $\frac{\varepsilon}{2}$-optimal, wenn x_n in $\mathcal{G}_i$ liegt.

Sei $(\mathcal{X} \cup \{*\}, \mathcal{A} \cup \{*\}, \Gamma', \rho', r')$ das $\mathcal{MM}$ mit:

$$\Gamma'(x) = \Gamma(x) \cup \{*\}$$

$$\rho'(x,a) = \begin{cases} \delta_x & , \text{ wenn } a = * \\ \rho(x,a) & , \text{ wenn } x \in \mathcal{X} - \bigcup_i \mathcal{G}_i, x \neq * \\ \delta_* & , \text{ wenn } x \in \mathcal{G}_i \cup \{*\} \end{cases}$$

$$r'(x,a,y) = \begin{cases} -1 & , \text{ wenn } x,y \in \mathcal{X} - \bigcup_i \mathcal{G}_i,\ a = * \\ r(x,a,y) & , \text{ wenn } x,y \in \mathcal{X} - \bigcup_i \mathcal{G}_i,\ a \neq * \\ 0 & , \text{ wenn } x = * \\ r(x,a,y) + V(y) & , \text{ wenn } y \in \bigcup_i \mathcal{G}_i \end{cases}$$

In diesem Modell ist die maximale mittlere Auszahlung $V'(x) = 1_{\cup_i \mathcal{G}_i} V(x)$, da nach Lemma 8.5 im ursprünglichen Modell jeder Plan $\bigcup_i \mathcal{G}_i$ mit Wahrscheinlichkeit 1 erreicht. In diesem Modell existiert ein stationärer Plan π', der $\bigcup_i \mathcal{G}_i$ mit Wahrscheinlichkeit 1 erreicht und dessen mittlere Auszahlung

$$\overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_{\pi'}[x] \geq V'(x) - \varepsilon = 1_{\bigcup_i \mathcal{G}_i} V(x) - \varepsilon,$$

ist (Satz 9.6). Der Plan π' benutzt die Aktion $*$ nicht. Denn würde er diese Aktion in einem Zustand y benutzen, so würde

$$\overline{\lim_{t\in T}} \int u_t \ \mathrm{d}\sigma_{\pi'}[y] \leq -1 + V'(y)$$

gelten, und der Plan π' wäre in y nicht ε-optimal.

Sei π der Markow-Plan mit:

$$\pi_j(x_0,\ldots,x_j) = \begin{cases} \pi_j^i(x_0,\ldots,x_j) & , \text{ wenn } x_j \in \mathcal{G}_i,\ i \text{ ungerade} \\ \pi_j^i(x_0,\ldots,x_j) & , \text{ wenn } x_j \in \mathcal{G}_i - \{\mathcal{G}_{i+1} \cup \mathcal{G}_{i-1}\} \\ \pi_j'(x_0,\ldots,x_j) & , \text{ wenn } x_j \notin \bigcup_i \mathcal{G}_i. \end{cases}$$

Der hier angegebene Plan π ist überall 2ε-optimal. ∎

14. Ein Beispiel.

Im Paragraph 13 wurde die Existenz eines ε-optimalen Markow-Plans gezeigt. Dabei wurde angenommen, daß der Zustandsraum endlich ist. In diesem Paragraph wird ein Beispiel eines $\mathcal{GM}$ angeben, in dem der Zustandsraum abzählbar ist, aber kein ε-optimaler Markow-Plan existiert. In diesem Beispiel ist das Übergangsgesetz ρ endlich additiv. Wie gezeigt wird, geht diese Eigenschaft des Übergangsgesetztes entscheidend in das Beispiel ein.

Im folgenden ist $(\mathcal{X}, \mathcal{A}, \Gamma, \rho, 1_g)$ ein $\mathcal{GM}$. Der Zustandsraum $\mathcal{X} = I\!N \cup \{b, g\}$ besteht aus der Menge der natürlichen Zahlen $I\!N$, einem guten Zustand g und einem schlechten Zustand b. $\mathcal{A}$ sei die Menge der Inhalte auf $\mathcal{X}$ und $\rho(x, a) = a$. 1_g bezeichnet die Indikatorfunktion von g und dient hier als Nutzenfunktion u.

Zur Beschreibung der Funktion Γ benötigen wir die folgenden Abkürzungen: δ_n is das auf $\{n\}$ konzentrierte Dirac-Maß. Für jede Menge $D \subset I\!N$ und jede natürliche Zahl z wird die Menge $\{n + z \,|\, n \in D\} \cap I\!N$ durch $D + n$ abgekürzt. δ bezeichne ein translationsinvariantes diffuses Maß (d.h. $\delta(D) = \delta(D + z)$ und $\delta(\{n\}) = 0$ für alle natürlichen Zahlen n, $\delta(I\!N) = 1$). Die Existenz eines solchen translationsinvarianten diffusen Maßes kann mit Hilfe des Hahn-Banach-Satzes (cf. Dunford-Schwartz (1957) S. 62) oder, wie D. Plachky mir mitteilte, mit dem Satz von Markov-Kakutani (cf. Dunford-Schwartz (1957) S. 456 und Übung VI 9.40) gezeigt werden.

$\Gamma(x)$ ist die Menge der erlaubten Aktionen mit

$$\begin{aligned}
\Gamma(n+1) &= \{\delta_n\} \text{ für jede natürliche Zahl } n, \\
\Gamma(b) &= \{\delta_b\}, \\
\Gamma(g) &= \{\delta_1\} \text{ und} \\
\Gamma(1) &= \left\{ (1 - \alpha - \alpha^2)\delta + \alpha\delta_g + \alpha^2\delta_b \,\middle|\, \alpha \in [0, 0.5] \right\}.
\end{aligned}$$

Zustände und Aktionen des $\mathcal{GM}$sind im Bild 6 dargestellt.

Das Beispiel kann als σ-additives Beispiel gesehen werden. Statt den Zustandsraum $\mathcal{X}$ zu betrachten, ist es für viele Überlegungen hilfreich, den

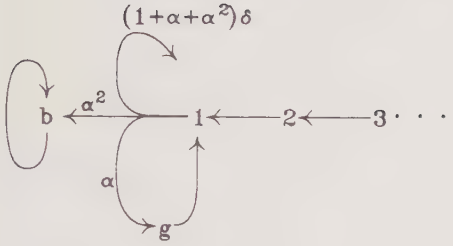

Bild 6

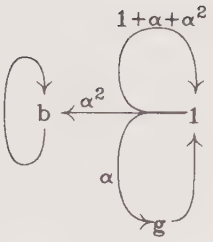

Bild 7

Zustandsraum $\mathcal{X}' = \{1, g, b\}$ zu untersuchen. Die erlaubten Aktionen im neuen Raum sind:

$$\Gamma(b) = \{\delta_b\},$$

$$\Gamma(g) = \{\delta_0\} \text{ und}$$

$$\Gamma(1) = \left\{ (1-\alpha-\alpha^2)\delta_0 + \alpha\delta_g + \alpha^2\delta_b \;\middle|\; \alpha \in [0, 0.5] \right\}.$$

Dabei dient 1_g wieder als Nutzenfunktion u. Der mittlere maximale Nutzen ist in diesem Beispiel für die Zustände 1 oder g 1 und 0 für den Startzustand b (Bild 7).

Für dieses neue Gambling Problem ist die Existenz von ε-optimalen Markow- und ε-optimaler 'tracking' Plänen bekannt (cf. Th.P. Hill (1979)). Dabei induziert der tracking Plan im neuen $\mathcal{GM}$ einen tracking Plan im ursprünglichen Modell, der die gleiche Auszahlung hat. Deshalb existiert im ursprünglichen Beispiel ein ε-optimaler tracking Plan. Es ist aber offen, ob ein ε-optimal tracking Plan generell existiert.

In dem Beispiel sind die Bedingungen von Lemma 12.2 erfüllt. Wähle dort $\mathcal{C} = I\!N \cup \{g\}$ und $\mathcal{B} = I\!N$. Ist daher δ nicht, wie hier angenommen, ein diffuses, sondern ein σ-additives Maß, so folgt die Existenz eines Markow-Plans aus Lemma 12.3. Die spezielle Wahl des Übergangsgesetzes geht daher

entscheidend in das Beispiel ein.

Wie der folgende Satz zeigt, ist der maximale mittlere Nutzen des $\mathcal{GM}$ $(\mathcal{X}, \mathcal{A}, \Gamma, \rho, 1_g)$ 1 für alle Zustände außer b.

Im Falle des Startzustands $x = 1$ werden wir die Stoppzeit

$$s_1(x_0, a_1, x_1 \ldots) := \min \left\{ n \in \mathbb{N} \,\middle|\, x_n = \left\{ \begin{array}{l} b \\ 1 \\ m,\ x_{n-1} \notin \{m+1, 1, g\}\ m > 1 \\ g,\ x_{n-1} \neq 1 \end{array} \right. \right\}$$

benutzen. s_1 stoppt in 1 und b oder wenn das System in nicht erlaubte Zustände übergeht. Mit t_x^n wird n-te Ankunft in x bezeichnet.

14.1 SATZ. $V = 1 - 1_b$.

BEWEIS: Sei $\varepsilon > 0$, seien ferner $(\varepsilon_i)_{i \geq 1}$ und $(\alpha_i)_{i \geq 1}$ zwei Folgen von positiven reellen Zahlen mit $\varepsilon_i < \varepsilon$, $\prod_i^\infty (1 - \varepsilon_i) > \sqrt{1 - \varepsilon}$ und $\alpha_i < \sqrt{1 - \varepsilon_i}^{-1} - 1$. Wähle einen Plan π, mit $\pi_i(p_i) := (1 - \alpha_j - \alpha_j^2)\delta + \alpha_j \delta_g + \alpha_j^2 \delta_b$, wenn $p_i \in (\mathcal{X} \times \mathcal{A})^i$ g j-mal enthält. Es wird gezeigt werden, daß der Nutzen dieses Plans π größer ist als $1 - \varepsilon$.

14.2 LEMMA. *Für jede Vorgeschichte $p \in (\mathbb{N} \cup \{g\} \times \mathcal{A})^i$ mit j Beobachtungen von g existiert eine finitäre Stoppzeit t mit*

$$\sigma_\pi[p](t_g < \infty) \geq \sigma_\pi[p](t_g \leq t) \geq 1 - \varepsilon_j.$$

BEWEIS: Sei $\sigma_\pi[p](t_g < \infty) = \sigma_\pi[p]\left(\bigcup_{n=0}^{\infty} (s_1^n \oplus 1 = t_g)\right) \geq \sqrt{1 - \varepsilon_j}$.

Aus Satz 5.2(*iii*) von R. Purves und W.D. Sudderth (1976) folgt die Existenz einer Stoppzeit t' mit $\sigma_\pi[p]\left(\bigcup_{n=0}^{t'} (s_1^n \oplus 1 = t_g)\right) \geq 1 - \varepsilon_j$. $t := \bigoplus_{n=0}^{t'} s_1^n \oplus 1$ ist die gesuchte Stoppzeit. ∎

FORTSETZUNG DES BEWEISES VON SATZ 14.1: Der Satz 14.1 folgt nun, da die Wahrscheinlichkeit, g unendlich oft zu durchlaufen, gerade die Wahrscheinlichkeit ist, g nach irgend einer Stoppzeit t zu erreichen [für diese finitäre Version des Satzes von Fatou siehe Theorem 9.4 von R. Purves und W.D. Sudderth (1973)]. ∎

Das Ziel der Untersuchung ist es, die Wahrscheinlichkeit, mit der g durch Markow-Pläne erreicht werden kann, zu berechnen.

14.3 LEMMA. *Ist für alle* $p \in (\mathcal{X} \times \mathcal{A})^*$ $\sigma_\pi[p](t_b < \infty) > c \quad c > 0$, *dann gilt*

$$\sigma_\pi[1](t_b < \infty) = 1.$$

14.4 LEMMA. *Es gilt* $\sigma_\pi[p1](x_1 \in I\!N, t_g < \infty) = 0$* *für alle* $p \in (\mathcal{X} \times \mathcal{A})^*$ *und jeden* ε*-optimalen Markow-Plan* π *mit* $\varepsilon < 1$.

BEWEIS: Wegen der Shiftinvarianz von δ gilt (l sei die Länge von p):

$$\sigma_\pi[p1](x_1 \in I\!N, t_b < \infty) = \int_{I\!N} \sigma_\pi[p1 \cdot n \cdots 1](t_b < \infty)\ \pi_l(p1)(dn)$$

$$\geq 0.25 \int_{I\!N} \sigma_\pi[p1 \cdot n \cdots 1](t_b < \infty)\ \delta(dn)$$

$$\geq 0.25 \int_{I\!N} \sigma_\pi[n \cdots 1](t_b < \infty)\ \delta(dn),$$

Im Fall, daß $\int_{I\!N} \sigma_\pi[n \cdots 1](t_b < \infty)\ \delta(dn) > 0$ ist, folgt $\sigma_\pi[1](t_b < \infty) = 1$ aus Lemma 14.3. Das widerspricht aber der ε-Optimalität von π. Deshalb gilt $\sup_\pi \sigma_\pi[p]\left(\bigcup_1^s t_1' \oplus 1 = t_g < \infty\right) = 0$. Durch transfinite Induktion über die Stoppzeit s folgt daraus $\int_{I\!N} \sigma_\pi[p_n 1](t_b < \infty)\ \delta(dn) = 0$. ∎

Die Nicht-Existenz eines ε-optimalen Markow-Plans ($\varepsilon < 1$) kann nun gezeigt werden.

14.5 THEOREM. *Es existiert kein* 0.09*-optimaler Markow-Plan.*

BEWEIS: Für jeden 0.09-optimalen Markow-Plan gilt $\pi_1(1)(\{b\}) = \pi_1(1)(\{g\})^2 < 0.09$. Lemma 14.1 impliziert: $\sigma_\pi[1](t_g < \infty) = \pi_1(1)(\{g\}) < 0.3$. Daher gilt $\sigma_\pi[1](x_n = g \text{ f.u.v. } n) \leq 0.3$, was aber der 0.09-Optimalität von π widerspricht. ∎

---

* *für* $p = (x_0, ..., x_l, a_l)$ *ist* $p1 := (x_0, ..., x_l, a_l, 1)$

15. Fatou Ungleichungen.

In dieser Arbeit werden oft Verallgemeinerungen des Lemmas von Fatou benutzt. Die hier benötigten Verallgemeinerungen wurden von R. Chen (1976) und Engelbert-Engelbert (1979) gezeigt. Sie gehen zurück auf eine Arbeit von W.D. Sudderth (1969). Hier sollen die Sätze wiedergegeben werden. Wegen der allgemeineren Form, in der die Sätze von R. Chen benötigt werden, werden die Beweise hier angegeben.

In diesem Paragraph sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $(\mathcal{A}_n)_n$ eine aufsteigende Folge von σ-Algebren auf $(\Omega, \mathcal{A})$ und $(X_n)_n$ eine zu $(\mathcal{A}_n)_n$ adaptierte Folge von Zufallsvariablen. Sei ferner T_∞ die Menge der μ-fast-sicher endlichen Stoppzeiten, T_e die Menge der endlichen Stoppzeiten und T_b die Menge der beschränkten Stoppzeiten.

15.a Fall fast sicher endlicher Stoppzeiten. Die beiden Sätze aus diesem Abschnitt gehen zurück auf eine Arbeit von W.D. Sudderth (1969). Diese Arbeit wurde erweitert durch Arbeiten von R. Chen (1976) und Engelbert-Engelbert (1979).

15.1 Satz (W.D. Sudderth (1969, Satz 1.1)).

(*i*) *Wenn alle Ausdrücke existieren, gilt:*

$$\overline{\lim_{t\in T_\infty}} \int X_t \ \mathrm{d}\mu \geq \int \overline{\lim_{n\in \mathbf{N}}} X_n \ \mathrm{d}\mu.$$

(*ii*) *Wenn alle Ausdrücke existieren, gilt:*

$$\underline{\lim_{t\in T_\infty}} \int X_t \ \mathrm{d}\mu \leq \int \underline{\lim_{n\in \mathbf{N}}} X_n \ \mathrm{d}\mu.$$

15.2 Satz (Engelbert-Engelbert (1979, Satz 2.1)).

(*i*) *Wenn* $\overline{\lim}_{t\in T_\infty} \int X_t \ \mathrm{d}\mu < \infty$ *ist und alle Ausdrücke existieren, so gilt:*

$$\overline{\lim_{t\in T_\infty}} \int X_t \ \mathrm{d}\mu = \int \overline{\lim_{n\in \mathbf{N}}} X_n \ \mathrm{d}\mu.$$

(*ii*) *Wenn* $\overline{\lim}_{t\in T_\infty} \int X_t \ \mathrm{d}\mu > -\infty$ *ist und alle Ausdrücke existieren, so gilt:*

$$\underline{\lim_{t\in T_\infty}} \int X_t \ \mathrm{d}\mu = \int \underline{\lim_{n\in \mathbf{N}}} X_n \ \mathrm{d}\mu.$$

15.b Fall endlicher Stoppzeiten. R. Chen (1976) zeigt, daß Integral und Limessuperior unter der Anahme der gleichgradigen Integierbarkeit vertauschbar sind. Dabei betrachtet er den Limessuperior über die beschränkten Stoppzeiten. Wir wollen diese Ergebnisse hier für eine beliebige Menge von Stoppzeiten erweitern.

15.3 SATZ (R. CHEN (1976, THEOREM 2.1)). *Sei $(X_n)_n$ eine Menge von zu $(\mathcal{A}_n)_n$ adaptierter Zufallsvariablen, so daß alle X_t $(t \in T_\infty)$ integrierbar sind und sei T eine Menge von Stoppzeiten $(T_b \subseteq T \subseteq T_\infty)$ mit folgenden Eigenschaften:*

mit t ist $t \vee j$ in T für alle $j \in \mathbb{N}$

und

mit t ist $t \wedge t'$ in T für alle $t' \in T_\infty$.

(i) *Ist $(X_n^-)_n$ eine gleichförmig integrierbare Menge von Zufallsvariablen, so gilt:*

$$\overline{\lim_{t\in T}} \int X_t \ \mathrm{d}\mu \geq \int \overline{\lim_{n\in \mathbb{N}}} X_n \ \mathrm{d}\mu.$$

(ii) *Ist $(X_n^-)_n$ eine gleichförmig integrierbare Menge von Zufallsvariablen, so gilt:*

$$\underline{\lim_{t\in T}} \int X_t \ \mathrm{d}\mu \leq \int \underline{\lim_{n\in \mathbb{N}}} X_n \ \mathrm{d}\mu.$$

BEWEIS: (ii) folgt direkt aus (i), weshalb nur (i) gezeigt wird. Die Integrierbarkeit von $\int \overline{\lim}_{n\in \mathbb{N}} X_n \ \mathrm{d}\mu$ folgt aus der folgenden Ungleichungskette. Dabei folgt die erste Ungleichung aus dem Lemma von Fatou und die letzte aus der gleichförmigen Integrierbarkeit von $(X_n^-)_n$.

$$\int \left(\overline{\lim_{n\in \mathbb{N}}} X_n\right)^- \ \mathrm{d}\mu = \int \underline{\lim_{n\in \mathbb{N}}} X_n^- \ \mathrm{d}\mu \leq \underline{\lim_{n\in \mathbb{N}}} \int X_n \ \mathrm{d}\mu$$

$$\leq \sup_n \int X_n^- \ \mathrm{d}\mu < \infty$$

Die Ungleichung wird nun mit Hilfe der Techniken aus Theorem 2.1(3) bei R. Chen (1976) gezeigt. Wegen Satz 15.2 gilt:

$$u := \overline{\lim_{t\in T_\infty}} \int X_t \ \mathrm{d}\mu = \int \overline{\lim_{n\in \mathbb{N}}} X_n \ \mathrm{d}\mu > -\infty.$$

Da $\sup_{t\le\tau,\tau\in T_\infty}\int X_t \ \mathrm{d}\mu \ge \overline{\lim_{t\in T_\infty}}\int X_t \ \mathrm{d}\mu$ für alle Stoppzeiten t in T_∞, kann für jedes $\varepsilon > 0$ und alle $t \in T$ ein $\tau_t \in T_\infty$ mit $\tau_t \ge t$ gefunden werden mit:

$$\int X_{\tau_t} \ \mathrm{d}\mu > u - \varepsilon.$$

Für jede Stoppzeit t und für alle $j \in \mathbb{N}$ sei $t_{t,j} := \min\{t \vee j, \tau_t\}$ und $X_{t_{t,j}} := 1_{[\tau_t < t\vee j]} X_{\tau_t}$. Es gilt $t_{t,j} \ge t$ und für alle j gilt $t_{t,j} \in T$. Aus dem Satz der monotonen Konvergenz folgt dann:

$$\int X^+_{t_{t,j}} \ \mathrm{d}\mu \to \int X^+_{\tau_t} \ \mathrm{d}\mu \text{ und } \int X^-_{t_{t,j}} \ \mathrm{d}\mu \to \int X^-_{\tau_t} \ \mathrm{d}\mu.$$

Deshalb existiert eine natürliche Zahl J_1 mit:

$$\int X_{t_{t,j}} \ \mathrm{d}\mu > u - 2\varepsilon \text{ für } j \ge J_1.$$

Da (u^-_n) gleichförmig integrierbar ist und $\sigma_\pi[x]([\tau_t < t \vee j]) \to 1 \ (j \to \infty)$ gilt, gibt es eine natürliche Zahl J_2 mit

$$\sup_k \int_{[\tau_t > t\vee j]} u^-_{t\vee k} \ \mathrm{d}\mu < \varepsilon \text{ für alle } j \ge J_2.$$

Bisher ist gezeigt, daß für alle $t \in T, \varepsilon > 0$ ein $j \ge \max\{J_1, J_2\}$ existiert mit:

$$\int X_{t_{t,j}} \ \mathrm{d}\mu > u - 3\varepsilon, \text{ da } t \text{ beliebig und } t_{t,j} \ge t \text{ folgt}$$

$$\overline{\lim_{t\in T}}\int X_t \ \mathrm{d}\mu > u - 3\varepsilon.$$

Mit $\varepsilon \to 0$ folgt das Lemma. ∎

15.4 Lemma (R. Chen (1976, Lemma 2.2)). *$(X_t)_{t\in T_b}$ ist genau dann gleichförmig integrierbar, wenn $(X_t)_{t\in T_\infty}$ gleichförmig integrierbar ist.*

15.5 Satz (R. Chen (1976, Theorem 2.2)). *Sei T eine Menge von Stoppzeiten $(T_b \subseteq T \subseteq T_\infty)$, für die eine Folge von Stoppzeiten $(t_n)_n$ in T mit $t_n \to \infty$ existiert.*

(i) *Sei $(X_t^+)_{t\in T_b}$ eine gleichförmig integrierbare Menge von Zufallsvariablen, so gilt:*

$$\overline{\lim_{t\in T}}\int X_t \ \mathrm{d}\mu \le \int \overline{\lim_{n\in\mathbb{N}}} X_n \ \mathrm{d}\mu \in \mathbb{R}.$$

(ii) *Sei $(X_t)_{t\in T_b}$ eine gleichförmig integrierbare Menge von Zufallsvariablen, so gilt:*

$$\underline{\lim}_{t\in T}\int X_t \ \ \mathrm{d}\mu \geq \int \underline{\lim}_{n\in N} X_n \ \ \mathrm{d}\mu \in \mathbb{R}.$$

BEWEIS: Es braucht nur (i) gezeigt zu werden. Die Integrierbarkeit von $\int \overline{\lim}_{n\in N} X_n \ \ \mathrm{d}\mu$ folgt aus der folgenden Ungleichungskette. Dabei folgt die erste Ungleichung aus Satz 15.4 (i). Wegen Lemma 15.4 folgt aus der gleichförmigen Integrierbarkeit von $(X_t^+)_{t\in T_b}$ die gleichförmig Integrierbarkeit von $(X_t^+)_{t\in T_\infty}$ und damit die letzte Ungleichung.

$$\int \left(\overline{\lim_{n\in N}} X_n\right)^+ \ \ \mathrm{d}\mu \ = \int \overline{\lim_{n\in N}} X_n^+ \ \ \mathrm{d}\mu \ \ \leq \overline{\lim_{t\in T}} \int X_t^+ \ \ \mathrm{d}\mu$$

$$\leq \sup_{t\in T_\infty} \int X_t^+ \ \ \mathrm{d}\mu \ < \infty$$

Ist t_n eine Folge von Stoppzeiten mit $t_n \to \infty$ und $\overline{\lim}_{t\in T} \int X_t \ \ \mathrm{d}\mu = \overline{\lim}_{n\in N} \int X_{t_n} \ \ \mathrm{d}\mu$, so gilt nach Lemma 1.2 in Chow-Robbins-Siegmund (1971):

$$\int \overline{\lim_{n\in N}} X_n \ \ \mathrm{d}\mu = \int \overline{\lim_{n\in N}} X_{t_n} \ \ \mathrm{d}\mu \geq \overline{\lim_{n\in N}} \int X_{t_n} \ \ \mathrm{d}\mu.$$

Aus der letzten Ungleichung folgt die Behauptung. ∎

15.c Fall endlich additiver Maße. Die in 15.a und 15.b angerührten Sätze sind auf die Theorie der Inhalte übertragbar. In diesem Abschnitt sei $(\mathcal{X}_n)_n$ eine Folge nicht leerer Mengen, σ_0 ein Wahrscheinlichkeits-Inhalt auf $\mathcal{X}_0$ und $\sigma := (\sigma_n)_n$ eine Folge von Übergangsinhalten von $\times_1^n \mathcal{X}_i$ nach $\mathcal{X}_{n+1}$. Wie in der Modellbeschreibung dargelegt, definiert die Folge σ ein endlich additives Maß auf $\times_1^\infty \mathcal{X}_i$. Dieses Maß wird ebenfalls mit σ bezeichnet.

15.6 SATZ (PURVES-SUDDERTH (1973, THEOREM 10.3)). *Ist g eine beschränkte, Borel-meßbare Funktion, $(X_n)_{n\in\mathbb{N}}$ sei die Folge der bedingten Erwartungswerte*

$$X_n(x_0 \ldots x_n, x_{n+1}..) := \int g(x_0 \ldots x_n, y_0, ..) \ \ \sigma[x_0, \ldots, x_n]\Big(d(y_0, \ldots)\Big),$$

dann gilt:

(i) *$(X_n)_n$ ist ein beschränktes Martingal,*

(*ii*) $(X_n)_n$ *konvergiert fast sicher,*

(*ii*) $\sigma\left\{h \;\middle|\; |\overline{\lim_{n\in\mathbb{N}}} X_n - g| > \varepsilon\right\} = 0 \quad \forall_\varepsilon.$

15.7 SATZ (PURVES-SUDDERTH (1973, THEOREM 10.4)). *Ist* $(X_n)_{n\in\mathbb{N}}$ *eine gleichförmig beschränkte Menge von Zufallsvariablen mit Struktur von* X_n *kleiner gleich n, dann gilt*

$$\overline{\lim_{t\in Te}} \int X_t \; \mathrm{d}\mu = \int \overline{\lim_{n\in\mathbb{N}}} X_n \; \mathrm{d}\mu.$$

16. Symbole.

$\mathcal{X}$	x	x_i	x_t	Zustandsraum, Zustände
$\mathfrak{X}$	$\mathfrak{A}$			σ-Algebren auf $\mathcal{X}$, $\mathcal{A}$
$\mathcal{A}$	a	a_i	a_t	Aktionenraum, Aktionen
$\mathcal{H}$	h	p		Menge der Pfade, Pfade, Teilpfade (Vorgeschichten)
px				Hintereinanderschaltung von Vorschichten
ρ	ρ_0			Bedingte Übergangswahrscheinlichkeit, Anfangsverteilung
Π	π	π_i		Menge der Pläne, Übergangswahrscheinlichkeit auf die Aktionen
$r(x_i, a_i, x_{i+1})$		u_t		Einschrittauszahlungsfunktion, Auszahlung bis Zeitpunkt t
V	$\underline{V}$			Maximale mittlere Auszahlung, max.mittlere Auszahlung bei dualer Nutzenfkt.
Γ	$\Gamma(x)$			Auswahlfunktion, Menge der erlaubten Aktionen
$\mathfrak{P}(.)$				Potenzmenge
$M(.,.)$				Menge der Wahrscheinlichkeitsinhalte
$P(.,.)$				Menge der Wahrscheinlichkeitsmaße
$t \oplus s$				Hintereinanderschaltung von Stoppzeiten
$t \wedge s$	$t \vee s$	$t_\mathcal{D}$	t_x	Minimum, Maximum von Stoppzeiten, Ankunftzeit
$[\alpha]$				Gaußklammerwert von α

17. Bibliographie.

[1] Bauer, H. (1974): *Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie;* De Gruyter, Berlin.

[2] Blackwell, D. (1964): *Memoryless strategies in finite-stage dynamic programming;* Ann.Math. Statist. 35, 863-865.

[3] Blackwell, D. (1965): *Discounted dynamic programming;* Ann. Math. Statist. 36, 226-235.

[4] Blackwell, D. (1965): *Positive dynamic programming;* Proc. of the 5 *th* Berkley Symp. on Math.Stat. and Prob. 1, 415-418.

[5] Blackwell, D. (1970): *On stationary policies;* J. R. Statist. Soc. 133, 33-37.

[6] Blackwell, D. (1976): *The stochastic process of borel gambling and dynamic programming;* Ann. Math. Statist. 4, 370-374.

[7] Bellman, R. (1957): *Dynamic Programming;* Princeton Univ. Press, Princeton.

[8] Bodewig, H.-H. (1979): *Über dynamische Optimierung mit endlich additiven Maßen;* Diplomarbeit, Universität Bonn.

[9] Chen, R. (1975): *A Note on the Dubins-Savage Utility of a Strategy;* Israel J.of Math. 21, 1-6.

[10] Chen, R. (1976): *Some Inequalities for Randomly Stopped Variables with Applicatines to Pointwise Convergence;* Zeit.f.Wahrsch.u.verw.Gebiete 36, 75-83.

[11] Chow, Y.S.; Robbins, H.; Siegmund, D. (1971): *Great Expectations: The Theory of Optimal Stopping;* Houghton Mifflin Company, Boston etc..

[12] Dawen, R.v. (1984): *Stationäre Politiken in Stochastischen Entscheigungsmodellen;* Dissertation, Bonn.

[13] Dubins, L.E.; Savage, L.J. (1965, 1976): *Inequalities for Stochastic Processes: How to Gamble if You Must;* Dover, New York.

[14] Dubins, L.E. (1974): *On Lebesque-like extensions of finitely additive measures;* Ann. Prob. 2, 456-463.

[15] Dunford, N.; Schwartz, J.T. (1958): *Linear Operators I;* Interscience Publishers, New York.

[16] Engelbert, A.; Engelberth, H.J. (1979): *Optimal Stopping and Almost Sure Convergence of Random Sequences;* Zeit.f.Wahrsch.u.verw.Gebiete 48, 309-325.

[17] Hill, Th.P. (1979): *On the existence of good Markov strategies;* Trans. Amer. Math. Soc. 247, 157-176.

[18] Hordijk, A. (1975): *Convergent dynamic programming;* Unveröffentlichtes Manuscript.

[19] Ornstein, D. (1969): *On the existence of stationary optimal strategies;* Proc.Amer. Math. Soc. 20, 563-569.

[20] Purves, R.; Sudderth, W.D. (1973): *Some finitely additive probability;* Technical Report no.220, University of Minnesota, Mineapolis.

[21] Purves, R.; Sudderth, W.D. (1976): *Some finitely additive probability;* Ann. Prob. 4, 259-276.

[22] Revuz, D. (1975): *Markov Chains;* Publ.Comp, Amsterdam.

[23] Rieder, U. (1976): *On optimal policies and marginals in dynamic programming;* J.appl.Prob. 13, 507-518.

[24] Schäl, M.; Sudderth, W.D. (1984): *Stationary Policies and Markov Policies in Borel Dynamic Programming;* preprint no.660; Universität Bonn.

[25] Strauch, R.E. (1966): *Negative dynamic programming;* Ann. Statist. 37, 871-890.

[26] Sudderth, W.D. (1969): *On the existence of good stationary strategies;* Trans. Amer. Math. Soc. 135, 399-415.

[27] Sudderth, W.D. (1971b): *A 'Fatou Equation' for Randomly Stopped Variables;* Ann. Math. Statist. 42, 2143-2146.

[28] Sudderth, W.D. (1981): *Gambling problems with a limit infimum payoff;* preprint no. 470; Universität Bonn.

[29] Wal, J.v.d. (1981): *Stochastic Dynamic Programming;* Math. Centre Tract 139, Mathematisch Centrum, Amsterdam.

Lebenslauf

Am 16. März 1952 wurde ich, Hans-Heinrich Bodewig, als Sohn des Prokuristen Paul Bodewig und seiner Ehefrau Clara-Maria, geb. Brauers, in Weeze, Kreis Kleve, geboren. Dort besuchte ich von 1958 bis 1963 die Katholische Volksschule und ging danach zum Alt- und Neusprachlichen Gymnasium der Abtei Marienstatt, Ww. Dort besuchte ich den altsprachlichen Zweig und legte im Juni 1972 das Abitur ab.

Im Wintersemester 1972/73 begann ich das Studium der Mathematik mit dem Nebenfach Informatik an der Rheinischen Friedrich-Wilhelms-Univeristät in Bonn. Im April 1975 legte ich dort die Vordiplomprüfung ab und die Diplomprüfung im April 1980.

Seit dem 15. Mai 1980 bin ich als wissenschaftlicher Angestellter am Institut für Angewandte Mathematik, Abteilung Wahrscheinlichkeitstheorie und Mathematische Statistik, im Sonderforschungsbereich 72 tätig.

Meine akademischen Lehrer waren die Herren Professoren S. Hildebrandt, P. Gabriel, G. Harder, O. Endler, C.-M. Ringel, W. Everling, H. Zima, M. Schäl und W. Vogel.